IDENTIFYING CONTEXT-SPECIFIC SYNTHETIC LETHAL MIRNA INHIBITION IN NON-SMALL CELL LUNG CANCER

APPROVED BY SUPERVISORY COMMITTEE

Alexander Pertsemlidis, Ph.D.

Michael White, Ph.D.

Pier Paolo Scaglioni, M.D.

Melanie Cobb, Ph.D.

David Corey, Ph.D

DEDICATION

To Jacek, Jadwiga and Brian Borkowski, and Alli Dryer.

IDENTIFYING CONTEXT-SPECIFIC SYNTHETIC LETHAL MIRNA INHIBITION IN NON-SMALL CELL LUNG CANCER

by

ROBERT JOHN BORKOWSKI

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2013

ACKNOWLEDGEMENTS

I would like to extend my most sincere thanks to everyone who has assisted and influenced me over the course of this project and my graduate career. First and foremost I send my thanks to my mentors, Dr. Alexander Pertsemlidis and Dr. Michael White, for their countless hours, their wisdom and insights and their motivation and insistence on my best efforts. While some may view multiple mentors as a challenge, I have relished the opportunity to channel their shared expertise into rewarding experiments, and to turn their more specialized expertise into a previously unimagined course for this project.

I thank my committee members, Drs. Pier Paolo Scaglioni, Melanie Cobb and David Corey for their insights and motivation throughout my training.

My work here was aided in part by many collaborators at UT Southwestern. I thank Bruce Posner and Gervaise Henry of the High Throughput Screening core facility for their assistance in the preparation and performance of the my inhibitor screen. I would also like to thank Dr. Michael Story and the Genomics Core of the Harold C. Simmons Comprehensive Cancer Center for their assistance in generating and analyzing transcriptional profiles. I also thank Chin-Rang Yang for sharing his RPPA data, and Ignacio Wistuba for sharing survival data from lung adenocarcinoma patients. Additionally, I would like to thank Dr. Rolf Brekken and Jason Toombs of the Brekken lab for their assistance in preparing orthotopic mouse xenograft experiments, as well as Michael Peyton and the Minna lab for their generosity of time, data and insights. Many

of the key observations in this project would not have been possible without their generosity.

I am incredibly lucky to have been surrounded by a variety of bright, insightful and compassionate colleagues throughout my graduate career. I thank all of the members of the Pertsemlidis and White labs that I have had the pleasure of working with throughout the years for their feedback, support, basic training, and encouragement. I send my special thanks to Christopher DeSevo, Clayton Boldt, Aileen Klein and Benjamin Shields for their support, insights, inspiration and, when necessary, commiseration. I am also especially grateful for my dialogues with my colleagues Dr. Brian Bodemann and Michael Torres through the process of preparing this document.

I would also like to thank the UTSW Cell and Molecular Biology Training Grant and all of its members for the financial support and broad and insightful conversations during my fellowship.

Last but absolutely not least, I would like to thank my family. I thank my parents, Jadwiga and Jacek Borkowski for the eternal support and encouragement, and for understanding why this took me so far away from home. I thank my brother, Brian Borkowski, for his unique combination of timely humor and willingness to critically listen to my work when I needed an ear. Finally, I send an incredible thanks to my best friend, my secret weapon and my wife, Alli Dryer. No one has ever inspired me, or helped me find my best self, like you have.

IDENTIFYING CONTEXT-SPECIFIC SYNTHETIC LETHAL MIRNA INHIBITION IN NON-SMALL CELL LUNG CANCER

Robert John Borkowski, B.A.

The University of Texas Southwestern Medical Center at Dallas, 2013

Supervising Professors: Alexander Pertsemlidis, Ph.D. and Michael A. White, Ph.D

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related fatalities in the US. This is due in part to a lack of highly effective therapies for advanced cases, and this is of special concern as most NSCLC cases are not diagnosed until they are in an advanced, later stage. Recent successes in developing genotypically-targeted therapies with potency only in a well-defined subpopulation of tumors suggests that identifying targeted therapies for additional common NSCLC genotypes will improve patient survival.

In this study I utilized a library of inhibitors to microRNAs, a class of post-transcriptional gene regulators, to identify novel synthetic lethal miRNA inhibition:molecular mechanism interactions in NSCLC. I accomplished this by screening a panel of 13 NSCLC and immortalized normal lung epithelium (HBEC) cell lines in two phases to identify miRNA inhibitors with selective toxicity in the NSCLC cell lines that were also benign in an HBEC cell line. Two inhibitors, the miR-92a and miR-1226* inhibitors, met these criteria. I then collected toxicity data in an expanded panel of 29 total cell lines. This expanded toxicity data was used to identify p53 loss as a molecular mechanism correlated with

sensitivity to the miR-92a and miR-1226* inhibitors in NSCLC cell lines. This was recapitulated by demonstrating sensitivity after knockdown of p53 in the previously resistant HBEC30KT cell line. I determined that the inhibitors were toxic in a very sequence-specific manner and that they down-regulated the miR-17~92 polycistron. Down-regulation of the polycistron was toxic in a context-specific manner, and the down-regulation of the miR-17~92 cluster in sensitive cell lines mimicked activation of a 1α, 25-dihydroxyvitamin D_3 response in NSCLC cell lines in a manner consistent with sensitivity to the miR-92a inhibitor.

The results of this investigation demonstrate that the screening approach utilized in this study was capable of identifying a synthetic lethal miRNA inhibition:molecular mechanism interaction, and that I was then able to use a genetically defined model of the mechanism to identify a relevant mechanism of action for the toxic inhibitors.

TABLE OF CONTENTS

PUBLICATIONS

Liqin Du, Maria C Subauste, Christopher DeSevo, Zhenze Zhao, Michael Baker, **Robert Borkowski**, Jeoffrey J Schageman, Rachel Greer, Chin-Rang Yang, Milind Suraokar, Ignacio I Wistuba, Adi F Gazdar, John D Minna, and Alexander Pertsemlidis. miR-337-3p and its targets STAT3 and RAP1A modulate taxane sensitivity in non-small cell lung cancers. *PLoS ONE* **7,** e39167 (2012).

Liqin Du, **Robert Borkowski**, Zhenze Zhao, Xiuye Ma, Xiaojie Yu, Alexander Pertsemlidis. A synthetic lethal screen for miRNA inhibitors regulating lung cancer cell survival and response to paclitaxel. *RNA Biology* **(IN REVISION)**

Benjamin B Shields, Chad V Pecot, Christa Nagel, Scott Purinton, Malia Potts, Ying Wang, Christina Ivans, Hyunseok Kim, **Robert Borkowski**, Jayantha Lea, Adi Gazdar, Keith A Baggerly, Anil K Sood, Michael A White. microRNA Induction of Fatal Differentiation Programs is a Common Vulnerability in Ovarian Cancer. **(IN PREPARATION)**

Robert Borkowski, Liqin Du, Chin-Rang Yang, Ignacio I Wistuba, Adi F Gazdar, John D Minna, Michael A White and Alexander Pertsemlidis. Inhibition of the miR-17~92 cluster by a miR-92a inhibitor induces 1α, 25-dihydroxyvitamin D3-like toxicity in non-small cell lung cancer. **(IN PREPARATION)**

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

2’-O-Me - 2'-O-methyl

2’-O-MOE - 2'-O-methoxyethyl

cDNA - complimentary DNA

CLL - Chronic Lymphocytic Leukemia

CNV - Copy Number Variant

DNA - Deoxyribonucleic Acid

EGFR - Epidermal Growth Factor Receptor

EMT - Epithelial to Mesenchymal Transition

ERK - Extracellular Signal-Regulated Kinase

IPA - Ingenuity Pathway Analysis

KRAS - V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog

LNA - Locked Nucleic Acid

LOH - Loss of Heterozygosity

miR - miRNA

miRNA - miRNA

NSCLC - Non-Small Cell Lung Cancer

PCR - Polymerase Chain Reaction

PI3K - Phosphatidylinositide 3-kinase

qRT-PCR - Quantitative Real-Time Polymerase Chain Reaction

RasGAP - Ras GTPase activating protein

RNA - Ribonucleic Acid

RNA-Seq - RNA Sequencing

RNAi - RNA Interference

RISC - RNA Induced Silencing Complex

RPKM - Reads Per Kilobase per Million

SNV - Single Nucleotide Variation

TCGA - The Cancer Genome Atlas

TKI - Tyrosine Kinase Inhibitor

VDR - Vitamin D Receptor

Chapter 1: microRNAs and cancer

1.1 Lung Cancer

1.1.1 Lung Cancer, A Historic Perspective

As a class, lung cancer describes tumors that originate in the lung. While lung cancer is the third most prevalent primary tumor site across all genders in the U.S. behind prostate and breast, it accounts for the most cancer related fatalities in the U.S., with a rate greater than prostate and breast combined [1]. In 2013, an estimated 228,000 new cases and 159,000 deaths will be attributed to lung cancer [2]. The prognosis for lung cancer is poor, with an overall five-year survival rate across all genders and races of approximately 16%, a small increase from 12% in the late 1970s [2].

However, if survival is categorized by the tumor's progression at the time of discovery, one sees that if the tumor is localized at discovery the five-year survival rate increases to 52%, if the tumor is regional the rate is 25% and if the tumor is in a distant stage the rate decreases to 4% [2]. This suggests two things: That early-stage lung cancers are much easier to treat successfully, and that the overwhelming majority of tumors are discovered in later, more advanced stages. Furthermore, this suggests that both early detection methods and treatment modalities for advanced stage tumors have yet to advance far enough to make significant impact on survival by either accelerating detection or increasing efficacy against late-stage lung cancers.

Tobacco use, in particular cigarette smoking, is understood to be the cause of the overwhelming majority of lung cancer incidents of any kind, accounting for 90% of cases [3]. The connection between tobacco usage and lung cancer is not a new one, as evidenced by its prominent appearance in the U.S. Surgeon General's report of 1964 [4] and the delayed but synchronized relationship between cigarette consumption and lung cancer incidents in American males shown (Figure 1.1).

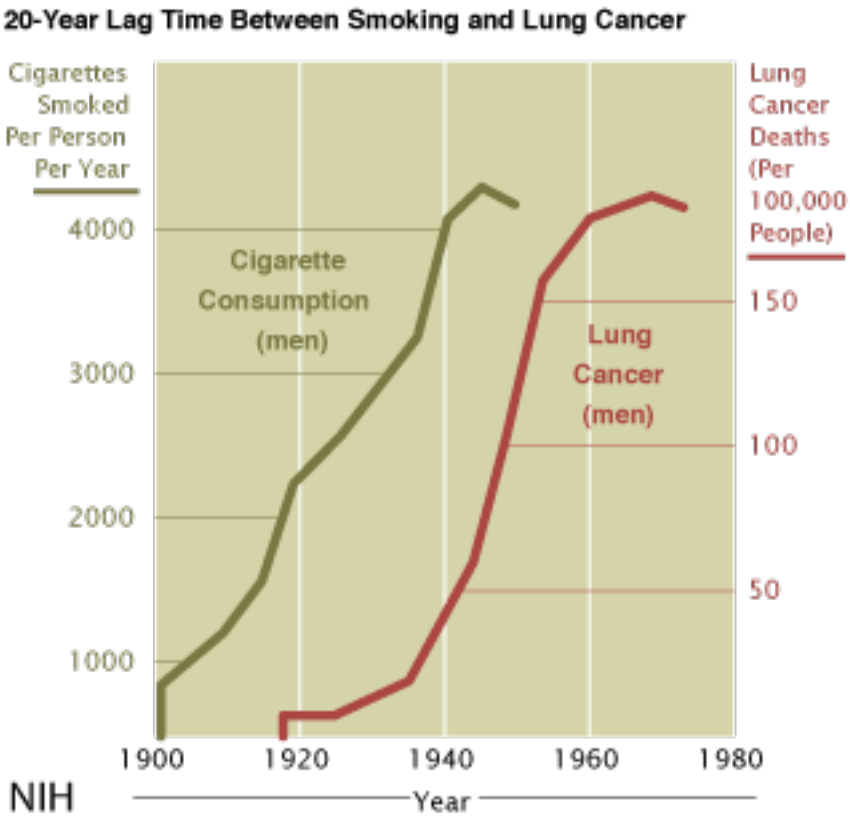


Figure 1.1 - Correlation between smoking and lung cancer in American males. Left vertical axis, in green, is cigarettes smoked per person per year. Right axis, in red, is lung cancer attributed fatalities in men per 100,000 people. Adapted from NIH published graphic, and http://en.wikipedia.org/wiki/File:Cancer_smoking_lung_cancer_correlation_from_NIH.svg.

1.1.2 Non-Small Cell Lung Cancer: Prevalence, mortality and first line therapy

Lung cancer is divided into two primary subgroups: Small cell (SCLC) and non-small cell lung cancer (NSCLC). SCLC is the far more clinically aggressive, with only 5-10% of patients surviving to five years, it accounts for a minority of cases at 15% [5]. The remaining 85% of cases are roughly identified as one of three histological subtypes of NSCLC, squamous cell carcinoma, large cell carcinoma and adenocarcinoma, with adenocarcinoma the most prevalent.

The same realities of progression and prognosis mentioned previously apply to NSCLC, with only 35% of NSCLC cases identified at Stage I or II, and a five-year overall survival in the single digits and approaching zero in advanced Stage III and Stage IV tumors [6]. This is due, in large part, to the efficacy of surgical resection on localized, early stage tumors; however, resectable tumors account for only 20% of all NSCLC tumors [5]. Even in cases where the tumor is successfully removed the five year survival rate is between 30-60% in patients who receive no follow-up therapy after surgery, with tumors formed beyond the thoracic cavity the major cause of recurrence-related deaths [7].

Most NSCLC patients receive some form of chemotherapy as part of their treatment, whether as an adjuvant therapy in the case of surgical resections, or as a first line response in non-resectable tumors [8-10]. The most common approach is to link a platinum-containing DNA crosslinker such as cisplatin or carboplatin with an anti-mitotic agent [8,11]. These anti-mitotic agents are compounds that interfere with the ability of

cells to proliferate and survive outside of normal regulatory signals. Taxanes (paclitaxel, docetaxel) and vinca alkaloids (vinorelbine, vincristine, vinblastine) target microtubules and either hyper-stabilize (taxanes) or destabilize (vinca alkaloids) microtubules. Nucleoside analogues, such as gemcitabine, and topoisomerase inhibitors like irinotecan prevent successful DNA replication. In 1995 the first definitive evidence of doublet (platinum-containing drug + anti-mitotic drug) therapy improving survival was published, and although the combinations of doublets tested in the clinic in vast, the improvement in survival is minimal, perhaps as low as 3-5% over five years [5,8].

1.1.3 The Genetic Landscape of NSCLC

At its origin cancer is essentially a genetic disorder, with lesions in a context-specific combination of pro-tumor oncogenes and anti-tumor tumor suppressor genes conferring a normal, regulated cell with the ability to divide freely and create the conditions for its proliferation and survival. After several decades of identifying the origins of various cancers by the cancer biology field, in 2000 Hanahan and Weinberg successfully distilled much of the work to date into the six hallmarks of cancer: Self-sufficiency in growth signal, insensitivity to anti-growth signals, evading apoptosis, limitless replicative potential, sustained angiogenesis and tissue evasion and metastasis [12]. Essentially, cancer is a stepwise progression from a normal, well-regulated cell that responds to growth and anti-growth signals to a heterogenous collection of cells that has acquired each of the aforementioned hallmarks [12].

Historically, cancer genetics consisted primarily of identifying genes whose activation promoted tumor growth(oncogenes), and genes whose impairment created an environment that permitted tumor growth (tumor suppressors). With regard to oncogenes, the 1976 report from Varmus and Bishop demonstrated that the DNA from an identified oncogene, *src*, was present in normal, untransformed cells [13]. This established that oncogenes exist in normal DNA as proto-oncogenes and that the gain of a hyper-activating modification to a proto-oncogene is necessary before tumor promoting actions such as aberrant proliferation and growth occur. In fact, many of the classical oncogenes identified in the past decades function in this manner. RAS, EGFR and PI3K require mutations and potentially amplification for oncogenic activity, and MYC is amplified in many tumor types [14-16]. Tumor suppressors repress or inhibit processes such as cell cycle progression and DNA damage response. The two classical tumor suppressor genes, Retinoblastoma (Rb) and p53, were shown to frequently lose their canonical functions by mutation or loss during tumor progression [17-19].

Identifying specific genes acting as the throttle (oncogenes) or brakes (tumor suppressors) in tumor formation and progression allowed for the identification of specific mechanisms involved in tumor formation and survival. The functions of the classical oncogenes and tumor suppressors such as KRAS, p53, Rb and MYC informed much of the structure and substance of the aforementioned “Hallmarks of cancer.” [12] While enumeration of the essential characteristics of cancer biology was a key step towards directed cancer research with the goal of developing more rational therapeutics, one

key area it did not fully address was the classification of tumors in a manner that could improve the granularity of research and improve outcomes.

Todd Golub and Eric Lander, in their seminal 1999 report on the utility of unbiased gene expression profiling in cancer class discovery and prediction state the following on the importance of classifying tumors in improving treatment [20]:

> The challenge of cancer treatment has been to target specific therapies to pathogenetically distinct tumor types, to maximize efficacy and minimize toxicity. . . Tumors with similar histopathological appearance can follow significantly different clinical courses and show different responses to therapy. . . For many more tumors, important subclasses are likely to exist but have yet to be defined by molecular markers. For example, prostate cancers of identical grade can have widely variable clinical courses, from indolence over decades to explosive growth causing rapid patient death. Cancer classification has been difficult in part because it has historically relied on specific biological insights, rather than systematic and unbiased approaches for recognizing tumor subtypes.

The publication of the human genome in 2004 was a key event enabling the sort of "systematic and unbiased approaches" to understanding tumor biology asked for in 1999 [21]. The new approaches to gaining insights into tumor biology took several different forms. Sequence-level information for the majority of transcripts in the genome allowed for higher coverage gene expression microarrays and gene-based classifiers for additional tumor types as well as for subtypes of histologically-similar tumors [22-25]. Comprehensive small interfering RNA (siRNA) libraries were created based on the published genome sequences and allowed investigators to perform genomewide forward genetic screens and identify phenomena such as synthetic lethal combinations of single gene depletions and chemosensitization in a model of NSCLC [26].

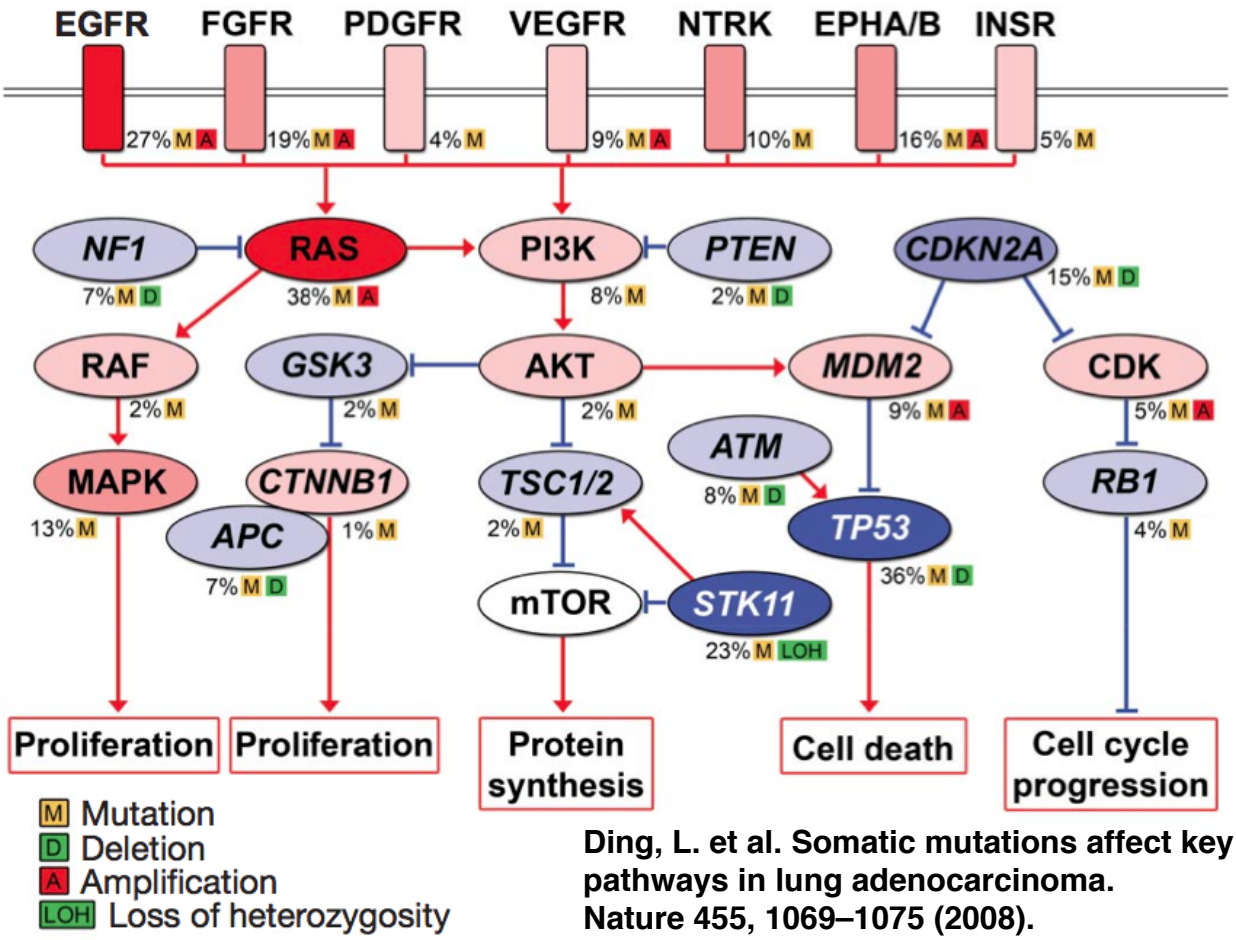
As cancer is at its essence a genetic reprogramming event that transforms normal cells into malignant and rapidly growing cells, the identification of somatic mutations in tumor types beyond the known oncogenes and tumor suppressors benefitted from the genomics era. Although truly unbiased approaches to mutation discovery were technically unfeasible due to the cost and resource intensity of full genome sequencing until the advent of “next generation” sequencing technologies in the late 2000s, several efforts proceeded to move beyond the suite of known cancer-related genes. One of the broadest efforts resulted in the Catalogue Of Somatic Mutations In Cancer (COSMIC) database and web site. Initially released in 2004 as a depot for 10,647 published mutations in HRAS, KRAS, NRAS and BRAF found in 57,444 tumors, COSMIC now (v64) combines published mutational information with tumor resequencing efforts and contains 913,166 mutations identified in 847,698 tumors across 24,394 genes [27-29].

Ding et al., in 2008, created an expanded map of lung adenocarcinoma’s mutational landscape [30]. Ding et al. surveyed the coding regions of 623 candidate genes including known oncogenes and tumor suppressors, protein kinases and genes in known regions of copy number aberration. They identified 1,013 non-synonymous mutations across the 188 primary lung adenocarcinoma tumors evaluated [30]. After integrating their identified somatic mutations with copy number and gene expression data for a superset of primary lung adenocarcinoma tumors they were able to identify several pathways that were enriched for genetic aberrations, including the MAPK pathway, where they saw at least one mutation in 132 of the tumors assessed [30]. Ding et al. demonstrated that thousands of of mutations in a panel of several hundred genes

could be reduced, with the assistance of gene expression and copy number analysis, to a handful of critical cellular pathways (Figure 1.2). Beyond providing a meaningful insight into the biology underlying lung adenocarcinoma Ding et al. demonstrated a biological basis for the identification of targeted therapeutic approaches to the treatment of lung adenocarcinomas: While there is substantial heterogeneity in genetic mutations or aberrations at the gene level, there is substantial homogeneity in the pathways these aberrations disrupt.

Studies such as Ding et al.'s identification of key genetic events in lung adenocarcinoma have initiated a gradual transition in tumor classification from histologically-based classification to classification based on oncogenic drivers and genetic signatures [31]. Pao and Girard, in their 2011 review, describe the traditional classification of lung cancers as binning into either Adenocarcinoma, Squamous or Large-cell NSCLC. After the importance of oncogenic drivers and oncogene addiction in NSCLC biology were understood, a shift began towards classifying tumors by known oncogenic drivers, with KRAS the first major driver identified in NSCLC [31-33]. Distinctions between *KRAS*-mutant and *KRAS*-wild-type NSCLC cases began in the late 1980s, and as genetic methods and databases improved other oncogenic drivers were identified including *EGFR, EML4-ALK, HER2, BRAF, MET, MAP2K1, AKT1* and *PIK3CA* [31]. Although there do appear to be diminishing returns at this point in time for identifying new oncogenic drivers in NSCLC and the prevalence of those new drivers, but the currently known ones do account for more than 50% of NSCLC cases [31].

As the technologies and expertise required to identify key genes, pathways and processes grew so did the depth of understanding of each of the “hallmarks” described in 2000. In addition to the fleshing out of the previously mentioned hallmarks, a 2011 revisit of the Hallmarks of Cancer by the same authors proposed four new points of emphasis in cancer biology, two emerging hallmarks, deregulating cellular energetics and avoiding immune detection, and two enabling characteristics, genome instability and mutation and tumor-promoting inflammation [34]. These new hallmarks are involved in the pathogenesis of at least some cancers, while the enabling characteristics facilitate the acquisition of the hallmarks necessary for tumor formation and survival [34]. These new traits emphasize a classic pattern observed in biomedical research: new questions always arise to fill the place of those sufficiently “answered,” and often new lines of investigation are driven by methodological advances.



Ding, L. et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455, 1069–1075 (2008).

Figure 1.2 - Genetic aberrations in key pathways in lung adenocarcinoma. Diagram of common genetic lesions in lung adenocarcinomas. Red coloring indicates gain of functions and hyperactivity in lung adenocarcinomas, blue represents loss of function or inhibitory interactions. Intensity of color is relative to the frequency of aberrations observed.

1.1.4 Targeted Therapeutic Approaches in NSCLC

Several genetically targeted therapeutic regimens have found success in the clinic, most notably for tumors with an EML4-ALK fusion protein, and for mutant *EGFR*. These targeted therapies are used in combination with a genetic test for the presence of a prognostic genetic lesion and therefore sacrifice broad applicability for high potency in the targeted subgroup. Two small molecule inhibitors of EGFR signaling, erlotinib and gefitinib, were recognized as EGFR antagonists, and early attempts at employment as a cancer therapy focused on treating cancers with high frequencies of EGFR-dependent tumors. While *in vitro* work showed high affinity for the compounds to the EGF receptor, early clinical trials did not demonstrate compelling response rates or improved overall survival [35]. However, these studies did not enroll patients based on EGFR status, and a later study noted that 13 of 14 patients who responded to gefitinib had a mutation in the tyrosine kinase domain of EGFR, and these mutations were not observed in non-responders [35,36]. This observation that mutational status of a tumor was an improved indicator of response to a targeted therapy than expression alone underscored the importance of identifying oncongenic “drivers” for relevant deployment of targeted therapies. While erlotinib and gefitinib were able, when properly targeted, to improve patient outcomes, they also highlighted one major pitfall of targeted therapeutic approaches: resistance. The median time to recurrence after EGFR TKI treatment was reported at 12 months, and analysis of resistant tumors showed that the majority of EGFR-TKI resistant tumors developed either a second, inactivating mutation in EGFR, or activation of a compensating receptor tyrosine kinase MET [37,38].

Therefore, while targeted therapeutic approaches have demonstrated efficacy in tumors harboring lesions predicting sensitivity, it is important to remember that tumors are rarely a homogenous cellular population and can adapt to selective pressures such as chemotherapy rapidly. A 2011 report from Navin et al. demonstrated that a single-cell sequencing method that generated copy number profiles of individual cells from solid breast tumors could allow for the derivation of the tumor's phylogeny [39,40]. The authors' analysis led them to conclude that tumors, at least the two they studied, adhered to a "punctuated clonal evolution" model of tumor evolution where tumor population branches are distinct from their common ancestor and without detectable intermediates [40]. This understanding of intratumor heterogeneity challenges the notion of treating a tumor as it is one homogenous object, although this phylogenetic analysis is insufficient for determining how different populations will respond to a common treatment.

Mass cytometry is a recently described method for sorting and profiling the state of single cells in a mixture of cells [41]. Mass cytometry is conceptually identical to flow cytometry, where cellular proteins are labeled with fluorescent antibodies and quantified one cell at a time; however, mass cytometry circumvents the issue of fluorescence emission spectrum overlap by labeling the antibodies with mass-based tags. This, in combination with a mass spectrometer as a detector, allows for multiplexing of 50 or more probes in one assay, versus 6 to 15 in classic flow cytometry [41,42]. This added depth allows for both highly granular classification of complex mixtures of cells (the hematopoetic system is a favorite model) and the concurrent interrogation of the state of

many intracellular signaling molecules. Of course, this depends on a prior knowledge of how to define subpopulations, but when that is possible very powerful investigations of the response of specific subpopulations to an insult or treatment in relation to the whole are possible [41,42]. Investigations of these sort on heterogeneous tumors that can be sub-classified have immediate relevance to drug discovery studies, and perhaps to identifying sensitive and resistant tumor components before or during treatment.

In light of this, understanding mechanisms of resistance to targeted therapies will allow for targeted therapeutic regimens that preemptively target resistance mechanisms and hopefully decrease recurrence or prolong the disease-free interval. Newly developed high content, single-cell assays such as single cell deep sequencing and mass cytometry should play a critical part in deconvoluting tumor heterogeneity and allow for the identification, just-in-case or just-in-time, of cells refractory to targeted therapeutic treatment.

1.2 microRNAs in Cancer

1.2.1 miRNAs emerge as important post-transcriptional regulators

In 1993 Lee et al. reported what, in retrospect, would be recognized as the first identified microRNA (miRNA) when they observed that in *C. elegans* the *lin-4* gene negatively regulated LIN-14 during normal larval development [43]. Lee et al. further observed that *lin-4* did not produce any protein, that the two RNA transcripts they identified were only 22 nt and 61 nt in length and that the 3' untranslated region (3' UTR) of LIN-14 contained a repeated sequence complimentary to *lin-4* [43]. The authors

observed that *lin-4* was a ~22 nucleotide (nt) RNA that post-transcriptionally down-regulates target genes through a sequence-driven interaction with the target gene's 3'UTR [44-49].

The next miRNA identified, *let-7*, was also in *C. elegans*. Loss of *let-7* was found to cause larval cell fates in adult cells, and ectopic *let-7* expression in larval cells induced adult fates [50]. Just as was the case with *lin-4*, *let-7* was a 21 nucleotide RNA that regulated genes with complimentarity sites in the 3' UTR [50]. Additionally, *let-7* was highly conserved and found in nearly all bilaterally symmetric animals, suggesting similar small-RNA mediated gene regulation may be present in other organisms [51]. In 2001 Hutvágner et al. demonstrated that *lin-4* and *let-7* small temporal RNAs were generated through the previously identified RNAi pathway, and that the enzyme DICER was required for both processes [52]. After identifying several highly evolutionarily conserved small RNAs and a conserved processing pathway there was enough information and motivation to begin identifying miRNAs en masse. Using the four-part criteria of expression of a ~22 nt mature RNA, location in an intergenic sequence, high DNA sequence conversation between *C. elegans* and *C. briggsae* and a ~65 nt stem-loop precursor, a group was able to identify 15 new small RNAs, which they now called microRNAs, in *C. elegans* and predicted that there may in fact be hundreds more of these genes present [47]. Additional concurrent cloning and informatics based approaches quickly brought the number of identified miRNAs to over 100 [45,46].

1.2.2 miRNA biogenesis and mechanism of action

The genesis and processing of miRNAs happens through a mechanism highly conserved from unicellular organisms through most plants and animals [49]. miRNAs begin as primary transcripts (pri-miRNAs) transcribed by RNA polymerase II in the nucleus [45,46,49,53]. miRNA genes are found in several configurations throughout the genome: some are found as solitary genes, some are part of a cluster of miRNAs that are transcribed together, and some miRNAs exist in or as introns between the coding regions of a host gene [49,54,55]. After transcription the miRNA forms a stem-loop structure ~60-70 nt in length called the pre-miRNA which is then cut out of the pri-miRNA by the nuclear RNase III Drosha [56,57]. The stem-loop pre-miRNAs are then exported from the nucleus into the cytoplasm for further processing in an Exportin-5 mediated process [58]. Once in the cytoplasm the RNase Dicer transforms the pre-miRNA into a ~22 nt double-stranded RNA [59]. During loading into a RNA-induced silencing complex with an Argonaute protein, typically Argonaute 2, as the catalytic component, typically the strand of the miRNA duplex with the weaker 5' pairing is loaded into the RISC [60-67].

Once a guide strand is loaded into the RISC the RNA-protein complex associates with, and typically decrease target mRNA levels through transcript destabilization or cleavage [68]. Determination of miRNA target genes is still a process full of vagaries, but occurrence of conserved miRNA target sites in the 3'UTR of target mRNAs is very consistent [69]. Perfect complimentarity between bases 2-7 as measured from the 5' end of the miRNA, known as the "seed" region, and the target mRNA was identified by

several groups as a consistently important element in accurate prediction of miRNA targets [69-73]. The ambiguity of miRNA target prediction beyond these common features led to the creation of many algorithms, all with unique perspectives on the salient features of miRNA:mRNA interaction. Although the degree of overlap between the predictions the various algorithms generate for a miRNA is very high, when there is disagreement it is often due to factors such as the weighting of sequence conservation of a target site, the treatment of the target nucleotide opposite the first miRNA nucleotide and thermodynamics were sources of much of the variation between different algorithms [69].

The appeal of creating a customized target prediction method speaks to a critical allure and frustration of today's miRNA biology. Namely, the ambiguity inherent in miRNA:mRNA interaction affords an investigator substantial freedom to find their narrative of interest in a list of possible biology. An unfortunate trend in reports on miRNA biology involves citing one of the aforementioned target prediction programs, almost universally TargetScan, as a "bioinformatic analysis" that consists of little more than identifying a gene of interest somewhere on the list of predicted targets to construct a single miRNA-single gene story. While assays exist to support these hypotheses, the nature of miRNA targeting suggests a one-to-many or many-to-many relationship between miRNAs and mRNAs, with a strong dependence on the stoichiometry of the miRNA and mRNAs of interest. In short, miRNA regulation of biological processes is highly dependent on the gene expression context in the system of interest.

1.2.3 miRNAs Involvement in Lung Cancer

The importance of miRNA dysregulation in cancer appeared soon after the identification of miRNAs as a key class of posttranscriptional regulators. Since Calin et al. identified frequent deletions of miR-15 and miR-16 in chronic lymphocyte leukemia in 2002 miRNAs have been implicated in the control of nearly all of the aforementioned "hallmarks of cancer" [74]. Additionally, in 2009 Kumar et al. demonstrated that *Dicer1*, the enzyme that trims a pre-miRNA hairpin into a mature duplex, is a haploinsufficient tumor suppressor, and that hemizygous loss of Dicer in these tumors resulted in general down-regulation of tumor miRNAs in a *KRAS*-driven mouse model of lung cancer [75]. Kumar et al. also observed a high frequency of hemizygous *DICER1* deletion in many common human tumor types. In tandem, these results demonstrate that miRNAs play a crucial role in the maintenance of a cellular homeostasis that keeps cell survival and proliferation under control. Figure 1.3, although not a comprehensive diagram of miRNA involvement in cancer, is meant to demonstrate the broad relevance of miRNA regulation in tumor initiation, progression, and survival. Additionally, the function of specific miRNAs is very context-dependent; a miRNA can function as a tumor suppressor in one context and as an oncomiR in another. Here I will concentrate on miRNAs relevant in the context of lung cancer.

Several large scale studies of human tumors and tumor-derived cell lines identified miR-21 as one of the most frequently up-regulated miRNAs in tumor biology. Impressively, one study observed miR-21 up-regulation versus normal adjacent tissue in

540 combined lung, breast, stomach, prostate, colon, and pancreatic tumors [76]. Additional studies identified its up-regulation in glioblastoma, head and neck carcinoma, ovarian cancer, B cell lymphoma, hepatocellular carcinoma and cervical carcinoma cell lines and tumors [77-84]. In lung cancer, miR-21's expression can be driven by the Ras/MAPK/ERK signaling pathway, and miR-21 then inhibits negative regulators of that signaling pathway as well as inhibiting several anti-apoptotic genes [77].

In an example of miRNA control of growth factor response in NSCLC, miR-128b has been shown to down-regulate the Epidermal Growth Factor (EGFR) receptor in NSCLC cell lines [85]. EGFR is a known hotspot for activating mutations in NSCLC, and LOH of miR-128b is shown to correlate with survival in a small sample of Gefitinib-treated East Asian NSCLC patients [85]. This is a somewhat confusing result as one would expect low EGFR, and therefore potentially high miR-128b levels to be a pro-survival condition; however, survival was measured post-Gefitinib treatment, and one could expect that only tumors with high levels of oncogenic EGFR would respond to the EGFR-targeted Gefitinib.

The miR-15/16 family of miRNAs has been shown to control the activity of several different cancer-related pathways in a context-dependent manner. miR-15/16 expression is inversely correlated with the expression of the anti-apoptotic B cell lymphoma 2 (BCL2) protein in CLL, a direct target of miR-15/16, and that introduction of miR-15/16 induced apoptosis in a CLL cell line [86]. In the NSCLC context both miR-15 and miR-16 are frequently down-regulated in tumors and cell lines, and that miR-15/16

was inversely correlated with CCND1, but not BCL2, expression in the tumors [87]. miR-15/16 was also observed to induce cell cycle arrest, not apoptosis, in a Rb-dependent manner in two NSCLC cell lines [87]. Importantly, the observation that miR-15/16 can exhibit anti-tumor effects by manipulating different pathways in different tumor types demonstrates the complexity and substantial layering of miRNA regulatory activity, especially in a case like this where a known target of miR-15/16, BCL2, does not appear to respond to miR-15/16 loss in NSCLC.

The miR-34 family of miRNAs is deleted or down-regulated in several tumor types. The miR-34a locus is deleted in many tumors, and miR-34b and miR-34c are both deleted in many breast carcinomas and significantly reduced in NSCLC [88-90]. Interestingly, the expression of miR-34 family members is highly correlated with p53 expression in tumors, and many p53-regulated genes are targets of the miR-34 family, including *BCL2, CDK4, CDK6* and *CCNE2* [88,91,92]. The loss of miR-34 protects tumors from cell cycle arrest and apoptosis, and introduction of miR-34 in lung adenocarcinoma cell lines was shown to induce apoptosis and G_1-arrest [93]. Interestingly, miR-34b over-expression sensitized a p53 wild-type NSCLC cell line, A549, to radiation, but not the p53-null NSCLC cell line H1299 [93].

The *let-7* miRNA family of miRNAs is known to regulate several key genes in lung cancer survival and progression. In 2005 Johnson et al. demonstrated that *let-7* could directly regulate RAS in *C. elegans* [94]. *KRAS*, one of the human homologues of the *C. elegans* RAS gene, is a known oncogene that is frequently amplified or up-regulated in

NSCLC. Johnson et al. observed down-regulation of *let-7* in lung tumors as well as an anti-correlation between *let-7* presence and Ras protein in lung tumor tissue [94]. A later study also identified SNPs in a *let-7* target site in the *KRAS* 3' UTR that increased lung cancer risk and KRAS expression in tumors, further establishing the importance of the *let-7*/*KRAS* relationship in NSCLC [95]. *let-7* was then shown to modulate protein abundance of HMGA2, and that introduction of a *let-7* resistant *HMGA2* was able to rescue *let-7* mediated decreases in proliferation [96]. HMGA2 is a transcription factor shown to be over-expressed in 90% of lung cancers and whose presence in the nucleus is associated with high cell proliferation and poor survival [96]. It is unclear how this observation of the anti-proliferative effects of *let-7* relates to the aforementioned *let-7* regulation of *KRAS* in NSCLC, although a later report on KRAS addiction in NSCLC observes that the cell lines employed here may not require KRAS activity for their survival [22].

The miR-200 family of miRNAs was established as a key regulator of the epithelial to mesenchymal transition (EMT) in many cancers, including ovarian, breast and lung cancer [97-100]. The transition of tumor cells from an epithelial to a mesenchymal program is of great interest as mesenchymal cells are more prone to invading surrounding tissues and generating metastatic lesions. These events almost universally correlate with poorer prognosis for patients with solid tumors in part because surgical resection becomes far more difficult as the number of tumors and their invasion of surrounding tissues increases. miR-200 is a negative regulator of EMT in lung cancer: decreased miR-200 was observed after EMT induction, and ectopic miR-200 expression

could abrogate EMT and invasion [97]. miR-200's reversible role in EMT regulation is due in large part to its targeting of E-cadherin transcriptional inhibitors ZEB1 and ZEB2. miR-200 down-regulation of ZEB1 and ZEB2 increases E-cadherin, an epithelial marker, and inhibition of miR-200 decreases E-cadherin and increases the expression of the mesenchymal marker vimentin [99]. Impressively, the miR-200/ZEB/EMT axis has been consistently shown to be the primary mechanism of miR-200 regulation of EMT across many solid, invasive tumors [97-101]. These findings provide a useful insight into the control of EMT and the involvement of miRNAs in that process. It additionally provides an avenue for addressing the challenge of halting the metastatic and invasive activity of an aggressive tumor. A combination of a EMT reversal through miR-200 restoration and induction of tumor cell death could improve outcomes in invasive tumors. One report showed that miR-200b could reduce metastasis and invasion in a murine model of hepatocellular carcinoma; however, portions of the E-cadherin promoter were methylated in the mesenchymal cells and a supplemental DNA methyltransferase was required in order to reduce invasion and metastasis [102]. Such a therapeutic approach may be beneficial as part of an adjuvant therapeutic approach to limit the formation of recurrent metastatic tumors, although an additional cytotoxic agent would still be required if metastases were already present.

Another area where miRNA have shown involvement in NSCLC is in chemosensitivity. miR-337-3p is a miRNA that is expressed highly in a fetal lung fibroblast cell line, detectable in HBEC cells, and generally lower/undetectable in NSCLC cell lines [103]. While over-expression of the miRNA alone was insufficient for toxicity in a NSCLC cell

line, miR-337-3p sensitized the cells to treatment with paclitaxel, a common first-line therapy in NSCLC, and caused G2/M arrest [103]. Identification of STAT3 as an *in silico* and *in vitro* target whose presence correlated with taxane sensitivity in NSCLC raises the interesting possibility of miRNA manipulation in a combinatorial treatment with an existing first-line therapeutic for treatment in a manner which requires less of both agents, although it remains unclear how potent the synergy between miR-337-3p and taxanes is in normal tissues.

These findings demonstrate that miRNAs are key components in pathways that maintain controlled cellular growth and localization NSCLC, such as miR-34 in p53 response and miR-200 in EMT. Additionally, other miRNAs, such as miR-21 and *let-7*, are up or down regulated in NSCLC cell lines or tumors versus normal controls in a manner that allows for dysregulation of cellular growth and evasion of cell death. As many of these miRNA-related cancer phenotypes involved increasing or decreasing mature miRNA levels relative to normal adjacent tissues, there exists the potential that miRNA “restoration” may be of therapeutic benefit in NSCLC tumors where miRNA dysregulation is involved.

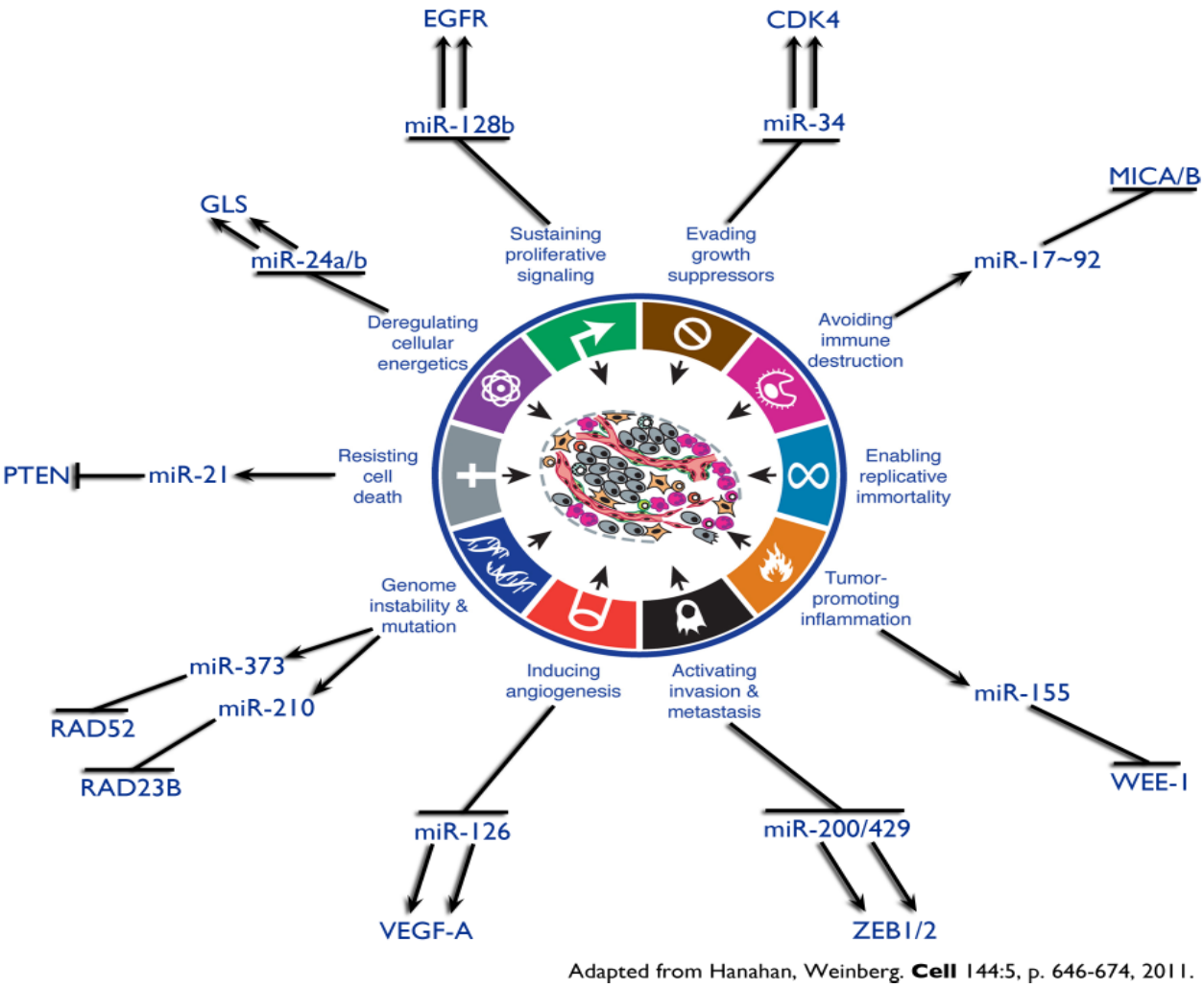


Figure 1.3 - miRNA involvement in the hallmarks of cancer. Arrows pointing to miRNAs indicate up-regulation in the given hallmark, bars indicate down-regulation. Corresponding arrows or bars pointing to genes in the outer circle indicate the result of that miRNA perturbation on gene expression.

1.2.4 miRNA manipulation as a therapeutic approach

miRNA-based therapeutics have been sought after for almost as long as miRNAs have been relevant to cancer biology. This is at least in part because gain- and loss-of function manipulations can be accomplished with similar approaches. The most clinically relevant approaches currently under development are based on oligonucleotide chemistries. Gain of tumor suppressor miRNA function can be achieved with chemically modified, double-stranded RNA oligonucleotides that act as miRNA mimics [104]. Single-stranded Locked Nucleic Acid (LNA) or modified RNA oligonucleotide inhibitors of oncomiRs can be employed to deplete miRNAs necessary for tumor survival [105-107].

There are two major success stories with regard to successful delivery of oligonucleotide-based miRNA mimics. Two tumor suppressors miR-34a and *let-7* have been developed as anti-tumor therapies in various mouse models of NSCLC. In 2010 Wiggins et. al. demonstrated successful intratumoral and systemic delivery of a mimic to miR-34a, a previously identified tumor suppressor miRNA in NSCLC, to A549 and H460 NSCLC xenograft tumors [104]. miR-34a. Wiggins et. al. utilized a neutral lipid emulsion to facilitate the cellular uptake of the the oligos into the xenograft tumors without observing any meaningful immune response in the mice after treatment [104]. A related group, Trang et. al., used a similar approach to deliver *let-7b* mimics by intratumoral injection to H460 xenograft tumors and observed an inhibition of tumor growth after treatment [108].

In a 2011 follow-up report Trang et. al. set out to determine whether the combination of double-stranded miRNA mimics and neutral lipid emulsion would be sufficient to impact tumor growth in native murine lung tumors. They observed a decrease in tumor burden in the lungs of LSL-*K-ras* G12D after 8 days of treatment with either an *let-7b* or miR-34a mimic [109]. They also delivered a benign miR-124 mimic by the same method and saw it delivered to liver, lung and heart tissues, as well as substantial amounts of miR-124 in blood; however, these measurements were made 10 minutes after injection and do not give any indication of the long-term stability or accumulation of the miRNA mimics in any tissues [109]. While Trang, Wiggins et. al. demonstrated some success in terms of delivering a therapeutically viable dose of a miRNA mimic to various mouse models of lung cancer, extremely high mimic dosages were required to inhibit tumor growth. The costs of treating human patients using methods similar to those employed here will be prohibitive until advances in delivery methods allow for higher efficiency in cellular uptake and more precise delivery of miRNA mimic-based therapies.

In addition to the therapeutic restoration of miRNA function mentioned above, there is increasing interest in the repression of oncogenic miRNAs whose presence or up-regulation promote tumor survival. Similar to therapeutic miRNA mimics the most promising approaches to *in vivo* miRNA inhibition involve oligonucleotide-based approaches. Anti-miRNA oligos are typically single-stranded oligos consisting of modified RNA or DNA bases [110]. Modifications such as LNA, 2'-O-Me, 2'-O-MOE serve to increase the stability, affinity and uptake of the oligos *in vivo* [110]. In 2005

Krutzfeldt et al. communicated the first successful case of oligo-based miRNA inhibition in mammals by tail vain injection of an “antagomir” to miR-16 in mice [105]. Antagomirs are 2’-O-Me - modified RNAs with a 3’ cholesterol tag to enhance delivery efficiency [105]. While Krutzfeldt et al. did not attempt to inhibit tumor growth or target any particular tissues they did demonstrate functional delivery of miR-16 and miR122 inhibitors to all the mouse tissues they assayed [105]. As with the miRNA mimics, the dosages of inhibitor used here were prohibitively expensive for a human therapy, although delivery was achieved without an additional packaging reagent [105].

In 2010 two reports, one each from Ma et al. and Anand et al., demonstrated the successful delivery of therapeutically relevant amounts of anti-miRNA compounds to tumors. Ma et al. used the previously described antagomir technology to inhibit tumor metastasis in a mouse mammary cancer model [106]. By delivering a miR-10b inhibitor to a highly metastatic breast cancer cell line xenograft beginning 48 hours post implantation Ma et al. were able to dramatically reduce the metastatic activity of the tumor [106]. However, the miR-10b inhibitor had no noticeable impact on growth of the primary tumor, only the degree of metastasis [106]. Whether this inhibition of metastasis would have any real impact on patient prognosis at the time of diagnosis when metastasis may already be widespread remains to be seen, although the potential for a potent antimetastatic agent in combination with surgical resection of a well defined primary tumor is promising.

In their 2010 report, Anand et al. targeted the miR-132/p120RasGAP/angiogenesis axis to decrease tumor burden in an orthotopic xenograft mouse model of human breast carcinoma. Anand et al. observed that miR-132 was up-regulated in human tumors versus normal endothelium, and that increasing expression of miR-132 in normal endothelium increased proliferation and tube-forming capacity [111]. They were then able to decelerate tumor growth and de-repress p120RasGAP using lipid nanoparticle-associated, 2'-O-Me modified anti-miR-132 oligos delivered to the orthotopic xenograft breast cancer tumors [111]. Using CD31 as a marker for vasculature, they showed a concurrent decrease in vasculature after anti-miR-132 delivery. However, tumor burden was decreased by only ~50% by the end of the study, and had not leveled off at that time, and it is unclear whether that is a reflection on the tumor's dependence on vascularization, or on the degree of vascular down-regulation achieved with the delivery system employed.

One interesting development in therapeutic miRNA inhibition in cancer treatment involves the targeted inhibition of families of miRNAs with a shared seed sequence. In 2011 Obad et al. explored the ability of 8-mer, fully LNA substituted oligos to inhibit miRNAs with a common seed sequence [112]. Obad et al. showed a successful proof of concept study with a "tiny LNA" inhibitor of miR-21 demonstrating de-repression of miR-21 targets by several methods only after delivery of a seed-targeted tiny LNA [112]. Obad et al. then used two different miRNA family-target tiny LNAs, one to the miR-221/222 family, and one to the let-7 family, and were able to demonstrate de-repression of luciferase constructs containing target sites for each family member

miRNA using a single family-specific oligo *in vitro* [112]. Next, they demonstrated that the tiny LNAs can be delivered to both orthotopic xenograft tumor models and normal tissues by delivering a tiny miR-21 inhibitor to the tumor and a tiny miR-122 inhibitor to the liver. However, they did not demonstrate any therapeutic benefit to miR-21 inhibition by a tiny LNA in the tumor model, and it is unclear how its potency relates to a classical, full-length miR-21 inhibitor.

While therapeutic manipulation of miRNAs in tumors has been heavily pursued over the past several years, with many reports of success in model organisms, several barriers remain between the current state of the art and a viable therapy. Currently, the synthesis of the modified oligos necessary for in vivo delivery is very expensive; one conservative dose of an LNA miRNA inhibitor for an average American male would cost approximately $150,000 for the oligo alone. There are two practical ways to reduce this figure to a realistic level: Decreasing the cost of the agent, and lowering the effective dose of the therapeutic agent. Much effort is being invested in the latter option, especially in terms of designing higher efficiency packaging molecules, although it will take time to achieve the orders of magnitude gains in efficiency necessary to make these interventions effective. An alternative approach worth added investigation involves finding synergistic interactions between gain or loss of a miRNA and an established antineoplastic compound, such as that reported between miR-337-3p and taxane sensitivity [103].

Chapter 2: Identifying conditionally toxic miRNA inhibition in NSCLC

2.1 Introduction

Lung cancer is the leading cause of cancer-related fatalities in the US, accounting for an estimated 159,480 deaths in 2013 [2]. Lung cancer can be divided into two major subtypes, small-cell and non-small-cell lung cancer (NSCLC), with the latter accounting for approximately 85% of all bronchiogenic carcinomas [2]. One subtype of NSCLC, adenocarcinoma, is the most common form of lung cancer, and has been shown to have a well-defined genetic background, with elevated mutation rates and copy number aberration in specific oncogenes and tumor suppressors such as *KRAS* and *TP53* [30]. The existence of defined genetic abnormalities in NSCLC has allowed for the development of targeted therapeutic approaches to NSCLC treatment, where the current five-year survival rate remains around 15%. Targeted therapies sacrifice breadth of treatable tumors for high efficacy in the presence of a specific biomarker for response to the treatment. In particular, therapies targeting tumors with mutations in the Epidermal Growth Factor Receptor (*EGFR*) gene or an *EML4/ALK* fusion gene have been clinically successful as first-line therapies [35,113-115]. However, only 25-35% of NSCLC tumors are targeted by the EGFR and *EML4/ALK* targeted therapies.

microRNAs (miRNAs) are a class of post-transcriptional regulators of gene expression: In a sequence-defined process mediated by the RNA Induced Silencing Complex (RISC), the highly conserved ~22 nucleotide RNAs down-regulate their target mRNAs

by associating with the 3' untranslated region (3' UTR) of their target mRNAs [49,68,69]. miRNA loci are found across the genome as either individual loci or as part of a polycistron, a group of miRNAs found on a shared primary transcript. Specific miRNAs have been implicated in many developmental processes, as well as cancer development, progression and chemosensitivity [16,77,90,103,116]. miRNAs can function as both tumor-promoters (oncomiRs) or tumor suppressors in various tumor types, and some miRNAs can play the role of either oncomiR or tumor-suppressor miRNA depending on the tumor type [117]. miRNAs are readily manipulated via genetic or oligonucleotide-based approaches both *in vitro* and *in vivo*, and miRNA gain or loss of function has been demonstrated to have substantial effects on tumor initiation and progression in *in vivo* tumor models [77,104,118]. Locked Nucleic Acid (LNA)-based oligonucleotides complementary to the sequence of a mature miRNA have demonstrated efficacy as therapeutic miRNA inhibitors [105,118,119] in both primate and murine *in vivo* cancer models due to their high target affinity and delivery to many tissue types without any packaging or carrier [105,118,119].

I was interested in using libraries of LNA-based miRNA inhibitors to identify synthetic lethal inhibitor:genotype interactions in NSCLC. Here I used a two phased screening approach to identify miRNA inhibitors with selective toxicity across a genetically diverse collection of NSCLC cell lines. I was then able to use the genetic diversity of the cell lines in tandem with mutational and transcriptional profiles of each cell line to identify a molecular mechanism, p53 loss, that correlated with miRNA inhibitor toxicity.

2.2 Results

2.2.1 A miRNA inhibitor screen converges on a specific selective toxicity in NSCLC cell lines.

In order to identify miRNA inhibitors with genotype-driven selective toxicity in NSCLC, I implemented a two phased miRNA inhibitor screen (Figure 2.2.1A). I employed a comprehensive library of miRNA inhibitors with 919 single-stranded competitive inhibitors targeting 870 of the known human miRNAs as of miRbase v.14. The DNA/ LNA hybrid inhibitors competitively bind to mature miRNAs and prevent miRNA loading into the RISC [120].

Two histologically similar yet genetically distinct NSCLC cell lines, H358 and H1993, were selected for comprehensive screening in Phase I. This was done to avoid differential response to an inhibitor based on the tumor's histological subtype. Both cell lines were derived from lung adenocarcinomas; however, only H358 harbors a known oncogenic *KRAS* allele [121]. I confirmed a differential response to *KRAS* depletion by introduced a small interfering RNA pool (siRNA) targeting *KRAS* by transient transfection into both cell lines. H358 was sensitive to the depletion of *KRAS*, while there was no difference in cell survival of H1993 after *KRAS* depletion versus control, confirming that while histologically similar the two cell lines can generate divergent responses to a specific perturbation (Figure 2.2.1A).

I observed a variety of cell viability responses to individual miRNA inhibitors between the two cell lines (Figure 2.2.1B). Specifically, I identified several miRNA inhibitors with high toxicity across both cell lines, as well as many other miRNAs with less potent but more selective toxicity in H358 versus H1993 (Figure 2.2.1B). 35 (3.79%) inhibitors decreased viability to below 50% of control in H358, and 11 (1.19%) in H1993. All 11 of the inhibitors decreasing viability below 50% in H1993 also decreased viability below 50% in H358. In order to identify inhibitors with selective, but not broad or idiosyncratic toxicity in NSCLC cell lines I selected 13 candidate inhibitors for additional profiling across an panel of 12 NSCLC and one human bronchial epithelial (HBEC) cell lines (Figure 2.2.1B, Table 2.2.1). Criteria for selection for the second phase included preferential toxicity in H358 versus H1993, differential expression between H358 and H1993, as well as over-representation of particular miRNA families in the toxic regions of the inhibitor response distribution (Figure 2.2.1B).

I observed three response patterns across the cell line panel: broad-spectrum toxicity across most or all cell lines, selective toxicity across subsets of the panel, or idiosyncratic toxicity in just one of the cell lines (Figure 2.2.1C). However, as my interest was in miRNA inhibitors with selective toxicity in specific classes of NSCLC I excluded broadly toxic inhibitors. Additionally, I excluded inhibitors with toxicity in the HBEC30KT cell line, as that indicated potential toxicity in non-transformed tissues and a diminished therapeutic window. Based on the selective toxicity of the hsa-miR-92a inhibitor, I chose the miR-92a inhibitor for profiling across an additional 13 cell lines, for a total of 28 cell lines (Figure 2.2.1D, Table 2.2.1). I observed a highly variable cell viability response to

the miR-92a inhibitor; there was no dichotomization of response to the miR-92a inhibitor into clearly defined sensitive and resistant groups (Figure 2.2.1D). This suggests that genetic factors predicting sensitivity to the miR-92a inhibitor may be continuously variable as well, versus a binary event such as the presence or absence of an activating/inactivating mutation.

miR-92a is a component of the miR-17~92 polycistron, a cluster of co-transcribed miRNAs that is frequently up-regulated in many solid tumors including those breast, colon, lung, pancreas, prostate, and stomach-derived tumors [76,122,123]. Up-regulation of this cluster of miRNAs was shown to promote proliferation in lung cancer cell lines [16]. miR-92a is also a member of the miR-25 seed family of miRNAs which includes miR-25, miR-32, miR-92a, miR-92b, miR-363 and miR-367 (Figure 2.2.2A). I did not observe toxicity of a similar magnitude to the miR-92a inhibitor from any inhibitors of either the miR-25 seed family of miRNAs (Figure 2.2.2B) or the miR-17~92 polycistron (Figure 2.2.2C), suggesting that the observed decrease in cell viability is highly specific to the activity of the miR-92a inhibitor.

The assay used in the primary screen, culturing relatively low numbers of cells in small-diameter wells with luminescent ATP quantification as a proxy for viability allowed us to identify differences in viability between cell lines as a function of treatment. However, because of the low cellular density, the assay was not a reliable discriminator between cytostatic and cytotoxic effects. I was therefore interested in identifying signals of cellular toxicity and apoptosis in sensitive cells. Using a luminescent assay for known

apoptotic markers activated Caspase 3/7 I observed a significant increase in apoptotic markers in the miR-92a sensitive cell line (H358) and not H1993 (Figure 2.2.1E). This suggests that the observed decrease in cell viability in response to the miR-92a inhibitor is due to a toxic response that induces apoptosis.

Interestingly, I observed that the screen appeared to converge on a very specific vulnerability in NSCLC. The two selectively toxic miRNA inhibitors, the miR-92a and miR-1226* inhibitors, appeared to converge on a specific vulnerability in NSCLC, as the toxicities of the two inhibitors were significantly correlated across the panel of cell lines (Figure 2.2.1F). Additionally, these inhibitors are identical by sequence at ten positions, including a “CAGGCC” motif at or near the 5’ end of each oligo (Figure 2.2.1G). This suggests that while I did not observe any highly specifically potent miRNA inhibitors in Phase I, I did identify a specific vulnerability in NSCLC cell lines that was perturbed by multiple miRNA inhibitors in the context of this screen.

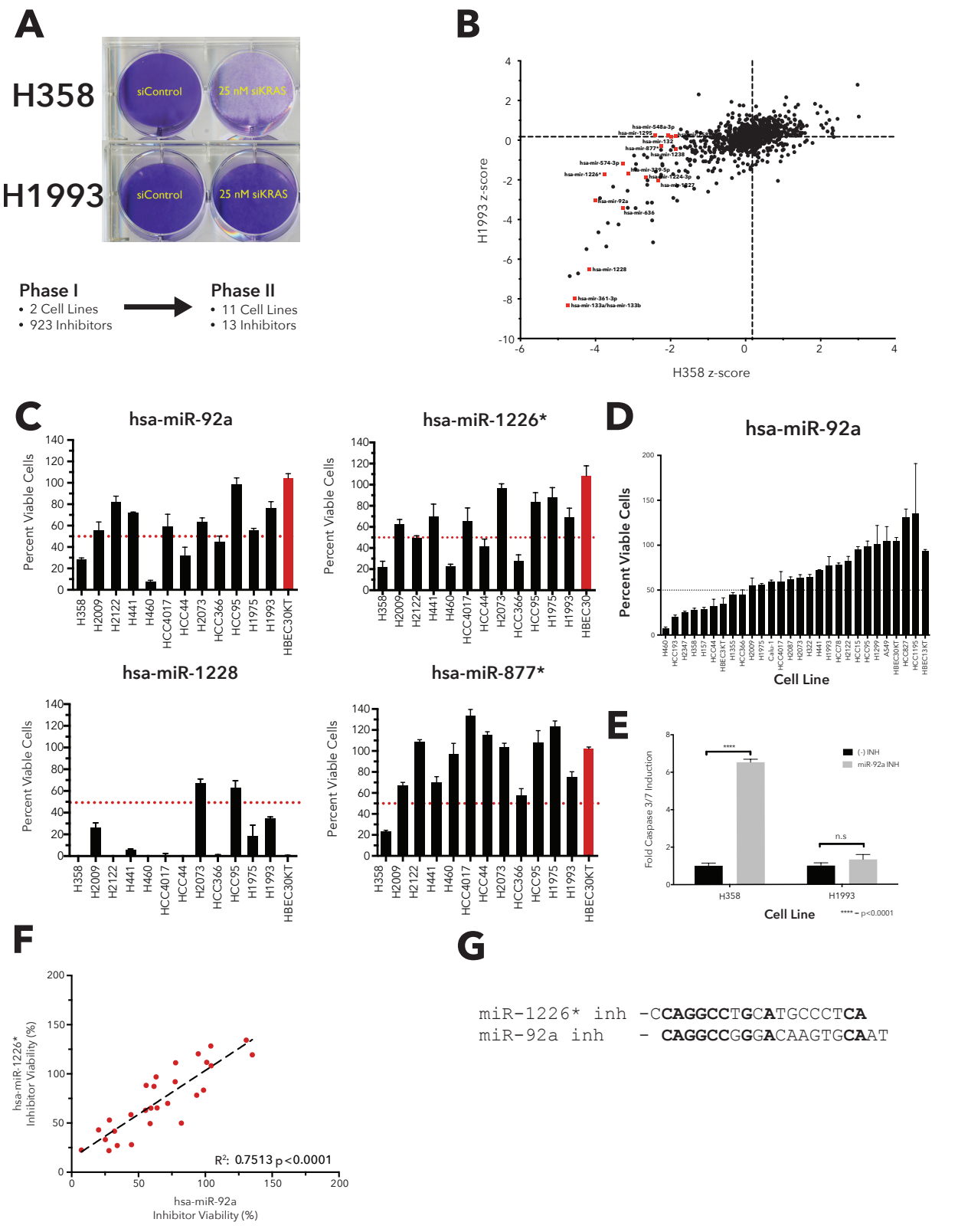

A
H358
H1993
siControl
25 nM siKRAS
siControl
25 nM siKRAS
Phase I
• 2 Cell Lines
• 923 Inhibitors
Phase II
• 11 Cell Lines
• 13 Inhibitors
B
H1993 z-score
H358 z-score
hsa-mir-548a-3p
hsa-mir-1225-3p
hsa-mir-132
hsa-mir-877*
hsa-mir-1238
hsa-mir-574-3p
hsa-mir-1226*
hsa-mir-1224-3p
hsa-mir-1827
hsa-mir-92a
hsa-mir-636
hsa-mir-1228
hsa-mir-361-3p
hsa-mir-133a/hsa-mir-133b
C
hsa-miR-92a
hsa-miR-1226*
hsa-miR-1228
hsa-miR-877*
Percent Viable Cells
H358 H2009 H2122 H441 H460 HCC4017 HCC44 H2073 HCC366 HCC95 H1975 H1993 HBEC30KT
D
hsa-miR-92a
Percent Viable Cells
Cell Line
E
Fold Caspase 3/7 Induction
(-) INH
miR-92a INH
****
n.s
H358
H1993
Cell Line
**** – p<0.0001
F
hsa-miR-1226* Inhibitor Viability (%)
hsa-miR-92a Inhibitor Viability (%)
R²: 0.7513 p<0.0001
G
miR-1226* inh -CCAGGCCTGCATGCCCTCA
miR-92a inh - CAGGCCGGGACAAGTGCAAT

Figure 2.2.1 - A miRNA inhibitor screen identifies several distinct toxicity patterns across NSCLC cell lines. A) Crystal violet staining in H358 and H1993 cells for cell viability 7 days post transfection with an siRNA pool to knock down *KRAS*. Schematic represents screen layout. **B)** Scatter plot of the 923 miRNA inhibitors in the inhibitor library used in the first phase of the screen. Each dot represents the Z-scores of the row-normalized cell viability in the two cell lines assayed, with H358 on the x-axis and H1993 on the y-axis. Dashed lines represent the median Z-score in each cell line. Labeled red dots are inhibitors selected for the second phase of the screen. **C)** Phase II results for inhibitors typical of the cell viability patterns observed in Phase I. miR-92a, miR-1228 and miR-877* results are percent viable cells relative to negative control as determined by CellTiter-GLO. The dashed red line indicates 50% viable cells. The red colored bars identify the representative normal cell line, HBEC30KT. Error bars represent one standard deviation. **D)** Same as in **C)**, with a total of 28 cell lines assayed for toxicity after transient transfection of 50 nM miR-92a inhibitor versus a control inhibitor. **E)** Selective activation of Caspase 3/7 72h post-transient transfection with the miR-92a inhibitor or control in H358 cells, versus H1993 cells. Results are the mean of 3 replicates. Error bars represent standard deviation. Significance was calculated by Student's two-tailed t-test. **F)** Correlation of cell viability in 28 NSCLC cell lines after miR-92a inhibitor or miR-1226* inhibitor treatment. Each red dot represents one cell line, dashed line represents a line of best fit. Correlation and p-value determined by Pearson correlation test. **G)** Sequence alignment of the miR-92a and miR-1226* inhibitors, from 5' (L) to 3' (R). Positions of sequence identity are bolded.

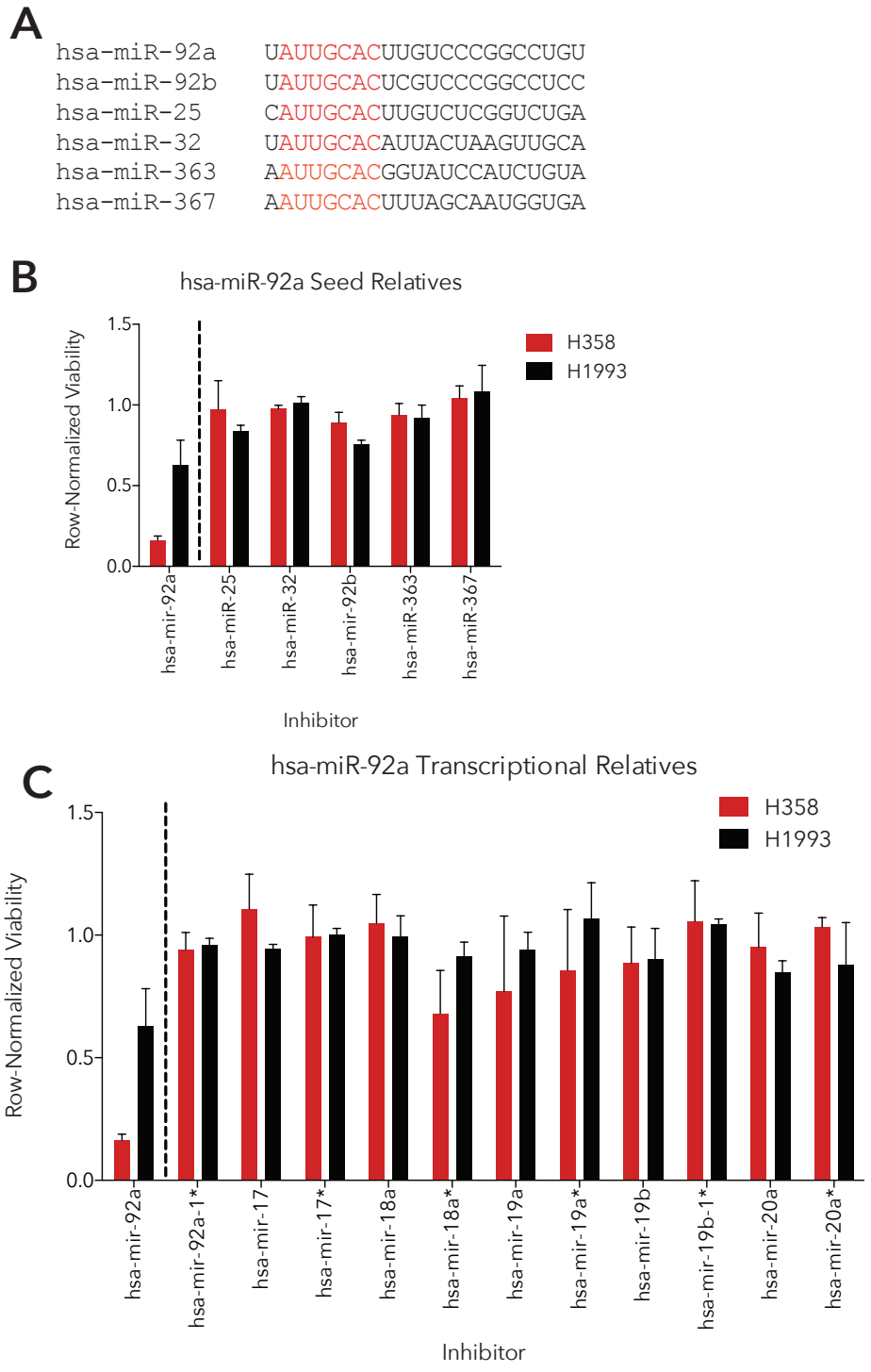


Figure 2.2.2 - Activity of miR-92a related inhibitors. A) Sequence comparison of the miR-25 miRNA family. The shared seed sequence is highlighted in red. **B)** Phase I toxicity of individual inhibitors of the miR-25 family of miRNAs. Values are the mean of 3 replicates in the labeled cell line, error bars are standard deviation. **C)** Phase I toxicity of individual inhibitors of the miR-17~92 cluster of miRNAs. Values are the mean of 3 replicates in the labeled cell line, error bars are standard deviation.

Cell Line	Tumor Type	Tumor Subtype	Age	Race	Gender	TP53 Status
NCI-H358	Lung	Adenocarcinoma	NA	Caucasian	M	Deletion
NCI-H2009	Lung	Adenocarcinoma	68	Caucasian	F	Mutant
NCI-H2122	Lung	Adenocarcinoma	46	Caucasian	F	Mutant
NCI-H441	Lung	Adenocarcinoma	33	Unknown	M	Mutant
NCI-H460	Lung	Large Cell	NA	Unknown	M	Wild-Type
HCC4017	Lung	Large Cell	62	Caucasian	F	Mutation
HCC44	Lung	Adenocarcinoma	54	Caucasian	F	Mutation
NCI-H2073	Lung	Adenocarcinoma	47	Caucasian	F	Mutation
HCC366	Lung	Adenosquamous	80	Unknown	F	Mutation
HCC95	Lung	Squamous	65	Caucasian	M	WT
NCI-H1975	Lung	Adenocarcinoma	NA	Unknown	F	Mutation
NCI-H1993	Lung	Adenocarcinoma	47	Caucasian	F	Mutation
A549	Lung	Adenocarcinoma	58	Caucasian	M	WT
Calu-1	Lung	Muco-epidermoid carcinoma	47	Caucasian	M	Deletion
Calu-6	Lung	Adenocarcinoma	61	Caucasian	F	Mutation
NCI-H1299	Lung	Large Cell Neuroendocrine	43	Caucasian	M	HD
NCI-H1355	Lung	Adenocarcinoma	53	Caucasian	M	Mutation
NCI-H157	Lung	Squamous	59	Caucasian	M	Mutation
NCI-H2087	Lung	Adenocarcinoma	69	Caucasian	M	Mutation
NCI-H2347	Lung	Adenocarcinoma	54	Caucasian	F	Mutation
NCI-H322	Lung	Adenocarcinoma	52	Caucasian	M	Mutation
HCC1195	Lung	Adenocarcinoma (mixed)	47	Black	M	Unknown
HCC15	Lung	Squamous	55	Black	M	Mutation
HCC193	Lung	Adenocarcinoma	71	Caucasian	F	Mutation
HCC78	Lung	Adenocarcinoma	55	Caucasian	M	Mutation
HCC827	Lung	Adenocarcinoma (BAC features)	38	Caucasian	F	Mutation
HBEC13-KT	NA	Immortalized Normal	65	Unknown	M	NA
HBEC3-KT	NA	Immortalized Normal	65	Unknown	F	NA
HBEC30-KT	NA	Immortalized Normal	62	Caucasian	F	NA

Table 2.2.1 - Cell Lines Used. Table naming cell lines employed in this study with selected annotation data including tumor type, patient age at tissue collection, race, gender and *TP53* mutational status.

2.2.2 miR-877* is an H358-idiosyncratically toxic miRNA that may be important in NSCLC.

In addition to the broadly toxic and selectively toxic inhibitors previously discussed, I identified one highly idiosyncratic miRNA inhibitor, the inhibitor for miR-877* (Figure 2.2.1C). While it reduced cell viability by 80% in H358, I could not identify any other cell lines in the expanded panel of 28 cell lines where viability was below 50% of control after miR-877* inhibitor treatment (Figure 2.2.3A). While this finding excluded miR-877* from follow-up as a candidate miRNA inhibitor for identifying genotype-specific vulnerabilities, I was intrigued by two observations. First, there appeared to be a correlation between expression of miR-877* and the toxicity of the miR-877* inhibitor (Figure 2.2.3B). Additionally, high miR-877* is correlated with poor prognosis in adjuvant-treated NSCLC patients, as determined by Ignatio Wistuba and collaborators (Figure 2.2.3C). Interestingly the differences in survival between miR-877* high and low patients was not significant in non-adjuvant treated patients (Figure 2.2.3C). These observations led me to assess the activity of this inhibitor in our NSCLC cell lines.

However, I made the surprising observation that, by qRT-PCR, miR-877* appeared to increase several orders of magnitude after miR-877* treatment using my standard protocol (Figure 2.2.4A, B). This was the case in three cell lines, and the effect was specific to the LNA-based Exiqon inhibitor. I used high-resolution acrylamide/TBE gel electrophoresis to qualitatively asses the presence of the miR-877* qPCR product in treated cells after a 36 hour treatment. The results of this assay confirmed that there

was in fact an increase in the miR-877* qPCR product in sensitive and resistant cells treated with the Exiqon miR-877* inhibitor (Figure 2.2.4C).

Still confused by this finding I repeated the standard reverse transcription and quantitative PCR process using either water or two concentrations of the resuspended miR-877* inhibitor as input. The purpose of this setup was to identify if the miR-877* inhibitor, as a short, single stranded LNA/DNA hybrid oligo, could act as a substrate for the probe oligo of the TaqMan miRNA assay. The TaqMan miRNA assay is a two-step process where a stem-loop hairpin primer is annealed to the target miRNA and reverse transcribed into a cDNA with the mature miRNA sequence and a universal primer sequence. The qPCR portion of the assay involves two primers, one targeted at the universal primer sequence and one miRNA-specific primer, as well as a detection probe in between the two primer sites. Surprisingly, I observed that, with or without the reverse transcriptase enzyme present for the RT step, the miR-877* inhibitor was amplified in a dose-dependent manner in the qPCR reaction (Figure 2.2.4D). Given the GC-rich 3' region of the mature miRNA, I believe that the affinity of the hydrolysis probe for the 5' end of the miRNA inhibitor is sufficient for polymerase landing and cleavage and release of the fluorophore from the miR-877* detection probe. Decreasing the affinity of the probe for the inhibitor by creating mismatched versions of the miR-877* inhibitor would allow for confirmation of this phenomenon. In total, this emphasizes that the small size of miRNAs can lead to concessions in detection assays that can result in counterintuitive results in the presence of miRNA mimics or miRNA-like molecules.

However, this is the only case of a LNA/DNA miRNA inhibitor acting as a cDNA template in a qRT-PCR reaction I have observed.

My last investigation of the miR-877* locus was an attempt to identify relevant target genes of miR-877*. This was more challenging than for a typical mIRNA, as this locus is not found in the gold standard TargetScan database, nor is there any literature on predicted targets of that particular miRNA. Therefore, I used the intersection of two different miRNA target prediction algorithms that contain results for this locus, miRmate and PicTar, to identify potentially relevant target genes. I identified three genes, *SLC12A9, RAD54L2* and *DEAF1* that were high-scoring predicted targets by both methods (Figure 2.2.5A). When I attempted to identify de-repression of target mRNAs by qRT-PCR, there was no coherent up-regulation of targets to greater than two-fold of the initial levels in the sensitive cell line H358; at 36 hours *SLC12A9* is up by approximately two-fold, but this is diminished after 72 hours (Figure 2.2.5B). Additionally, in the set of identified mRNA targets of the miR-877* inhibitor, *SLC12A9* was one of the most abundant across the panel of cell lines and the other two targets assayed, *RAD54L2* and *DEAF1* were of relatively low abundance, suggesting that there may be some range to de-repress these mRNAs in the proper context, or that these genes are in a repressed or inactivated state in NSCLC (Figure 2.2.5C).

Without sensitivity to the miR-877* inhibitor in more than one NSCLC cell line I did not have the power to generate other leads by interrogation of genetic profiles of the cell lines as to the biological impact of this potentially interesting but definitely persnickety

miRNA inhibitor. While evidence exists in the form of patient survival data for the importance of this miRNA in NSCLC I decided to focus my energy on other miRNA inhibitors that afforded us more dynamic range to identify genetic markers correlated with toxicity that would lead me closer to understanding the biology of the miRNA's activity. However, we observed a significant difference in survival between low-miR-877* and high-miR-877* expressors in NSCLC, but this was only observed in adjuvant treated patients (Figure 2.2.3C). This suggests that while I observed miR-877* exhibiting toxicity on its own in H358, its relevance in the broader NSCLC context may have to do with tumor response to adjuvant chemotherapy. Further work will be required to identify the full relevance of miR-877* in NSCLC and its impact on survival in adjuvant-treated tumors.

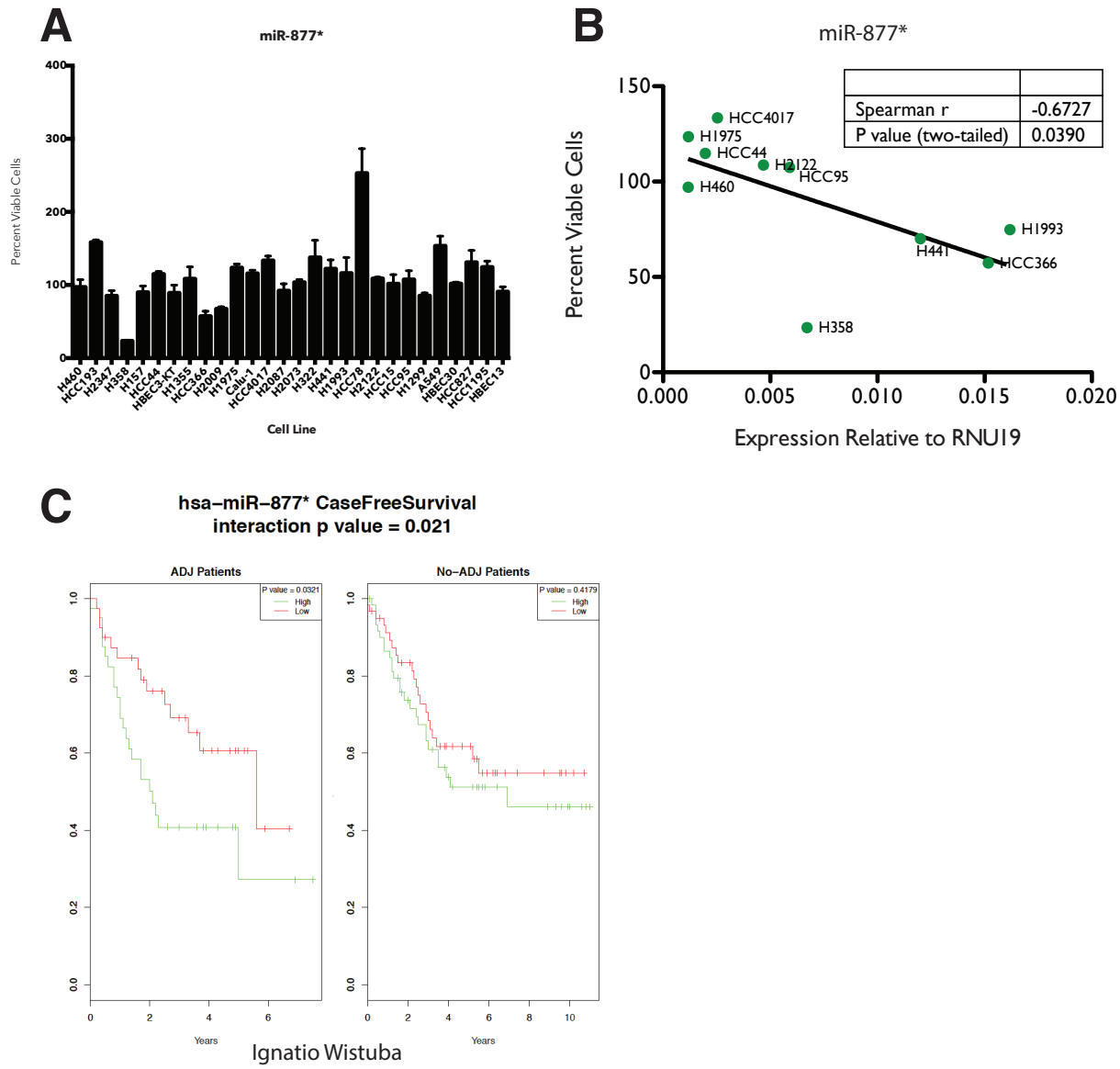


Figure 2.2.3 - miR-877* in NSCLC. A) miR-877* inhibitor toxicity across NSCLC cell lines. **B)** Correlation between miR-877* abundance and toxicity in NSCLC cell lines. Correlation determined by Spearman rank correlation. **C)** Kaplan-Meier survival analysis for miR-877* in lung adenocarcinoma. Left panel is adjuvant treated patients, right panel is non-adjuvant treated.

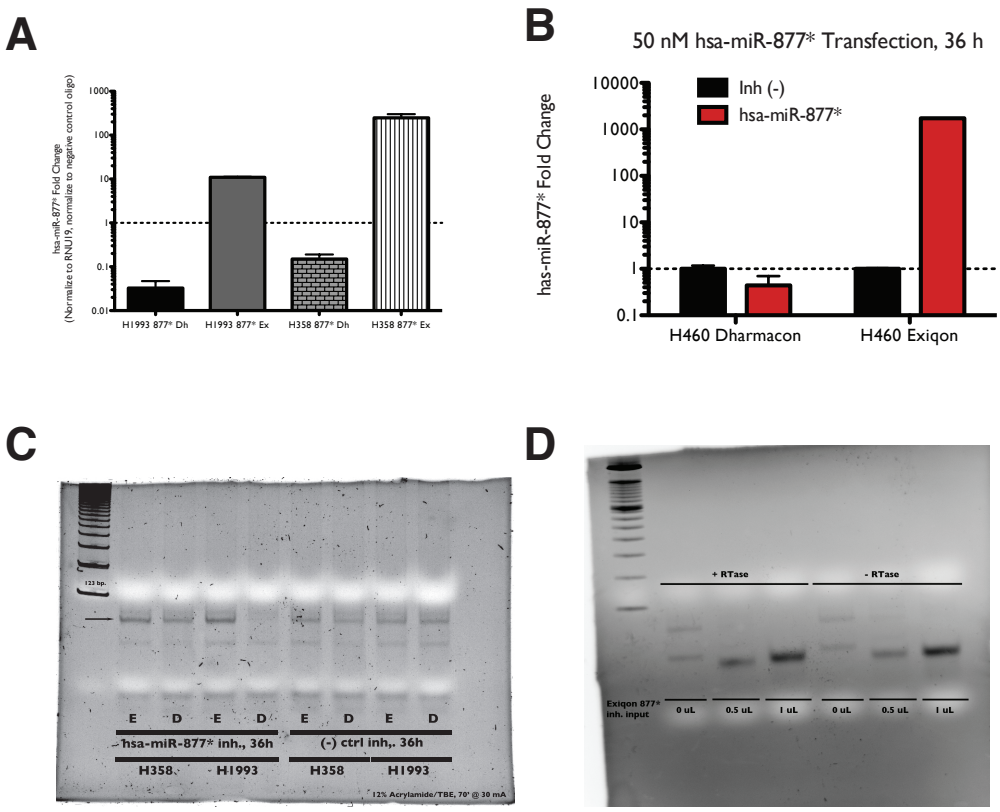


Figure 2.2.4 - Functional analysis of the miR-877* inhibitor is hindered by qRT-PCR detection of the miRNA inhibitor. A) qRT-PCR for miR-877* after treatment with LNA or RNA miR-877* inhibitors. Values are fold change relative to a chemistry-matched negative control oligo. Bars = mean, error = s.d., n=3. **B)** Short-term miR-877* treatment in a sensitive cell line. qRT-PCR data generated and plotted as in **A**. Black bars are 50 nM negative control treated, red bars are 50 nM miR-877* treated. **C)** High-resolution acrylamide gel of miR-877* qPCR products from **A**. E = Exiqon inhibitor, D = Dharmacon inhibitor. Arrow marks band of interest **D)** Acrylamide gel as in **C**. RTase = Reverse Transcriptase.

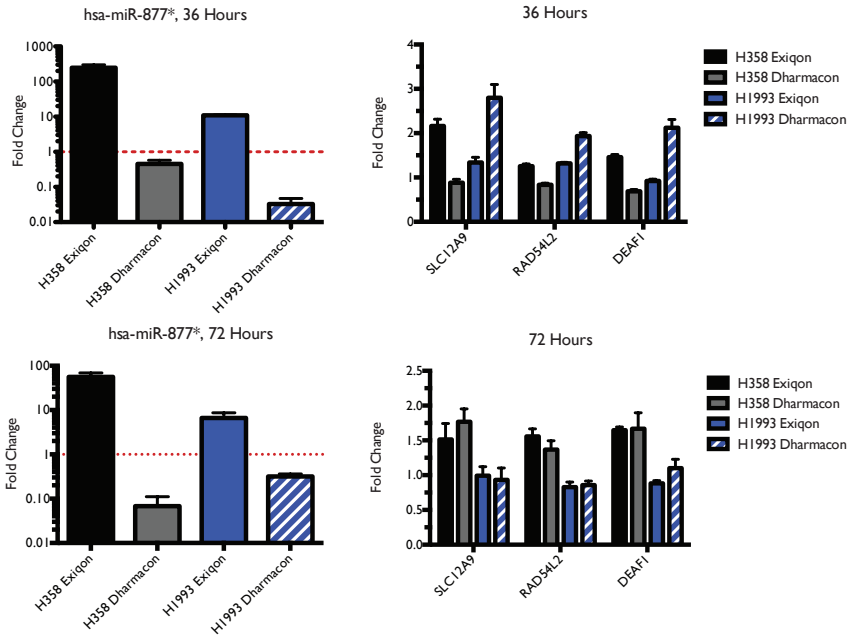
A

```
3'- GACCCUCCUCCCUCUUCUCCU -5' [miRNA]
    || |||||||| :|||||||||
5'- CUUGGAGGAGGAGGAAGAGGA -3' [SLC12A9]

3'- GACCCUCCUCC--CUCUUCUCCU -5' [miRNA]
    |||||:||| |  |||||||||:
5'- CUGGGGGGACGUUGAGAAGAGGG -3' [DEAF1]

3'- GACCCUCCUCCCU----CUUCUCCU -5' [miRNA]
    :||| ||||||||     |||||||||
5'- UUGGCAGGAGGGAAAAGGAAGAGGA -3' [RAD54L2]
```

B



C

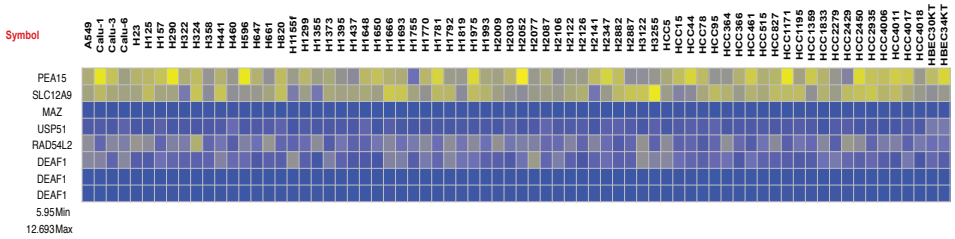


Figure 2.2.5 - Functional analysis of the miR-877* inhibitor predicted targets A) Target site alignments between miR-877* and predicted targets. **B)** qRT-PCR at 36 hours (top) or 72 hours (bottom) for predicted miR-877* targets. **C)** Heat map of predicted miR-877* target gene expression in NSCLC cell lines. Scale, from lowest to highest expression, is from blue to yellow.

2.2.3 p53 levels anti-correlate with miR-92a inhibitor toxicity, and loss of p53 sensitizes a resistant cell line to the miR-92a inhibitor.

Therapeutic approaches in NSCLC that involve genotype-based targeting have shown promising returns**,** particularly for those using EGFR or EML-ALK4 status as a biomarker for treatment [114,124]. I therefore wanted to identify a genetic predictor of the miR-92a inhibitor's toxicity (Figure 2.2.1D) by identifying correlations between the inhibitor's toxicity and genetic lesions in the NSCLC cell lines.

After interrogating the mutational status and gene expression of known oncogenes and tumor suppressors in NSCLC for correlates with miR-92a inhibitor toxicity I observed that *TP53* mRNA levels were anti-correlated with the toxicity of the miR-92a inhibitor: cell lines with low *TP53* had more robust toxic responses to the inhibitor (Figure 2.2.6A). Neither cell line used in the first phase of the screen contained an intact, wild-type *TP53* locus (Table 2.2.1). H358, the sensitive cell line from the first phase of the screen, had the lowest observed *TP53* expression of any of the cell lines in the panel (Figure 2.2.6A).

Recent studies have demonstrated that mutant p53 alleles can create either gain or loss of function reprogramming of p53 protein [125,126]**.** Using reverse phase protein arrays to quantify p53 protein abundance in the cell lines, I interrogated the correlation with miR-92a toxicity (Figure 2.2.6B) and observed that the correlation between toxicity and p53 protein was nearly identical to that observed in the mRNA data, and the relative placement of H358 and H1993 p53/*TP53* abundance in the panel of cell lines was

consistent between the data sets (Figures 2.2.6A, 2.2.6B). Through immunoblotting for p53 in the cell lines used in the first phase of the screen after transient transfection of either a control miRNA inhibitor or the miR-92a inhibitor I did not detect p53 in H358 and found abundant p53 in H1993, indicating that a total loss of p53 may be a miR-92a inhibitor-sensitizing event in NSCLC (Figure 2.2.6C). As p53 is an important responder to cellular stress, I decided to assess what, if any, effect the LNA inhibitors had on cellular levels of p53 [127-129]. After transient transfection of the miR-92a inhibitor or a negative control, I did not observe any chemistry or sequence dependent changes in p53 abundance in either a sensitive or resistant cell line 72 hours post treatment (Figure 2.2.6C).

After identifying that low *TP53* mRNA and low p53 protein correlate with miR-92a inhibitor toxicity in NSCLC, I employed a genetically controlled model of p53 loss in bronchial epithelium to address the necessity or sufficiency of p53 loss in this cellular context with regards to the inhibitor's toxicity. A CDK4, hTERT immortalized human bronchial epithelial cell line, HBEC30KT, was modified [130] to constitutively produce short hairpin RNA (shRNAs) to *TP53*, creating a knockdown of p53. The immortalized an untransformed HBEC30KT cell line was used as a representative normal lung tissue cell line in Phase II of the screen, and, as expected, did not respond to the miR-92a inhibitor (Figure 2.2.1C). The knockdown of p53 in HBEC30KT was sufficient to render the previously resistant cell line sensitive to the miR-92a inhibitor, confirming our previous observation that a loss of p53 protein is a miR-92a inhibitor-sensitizing event in tissues derived from the lung epithelium (Figure 2.2.6D).

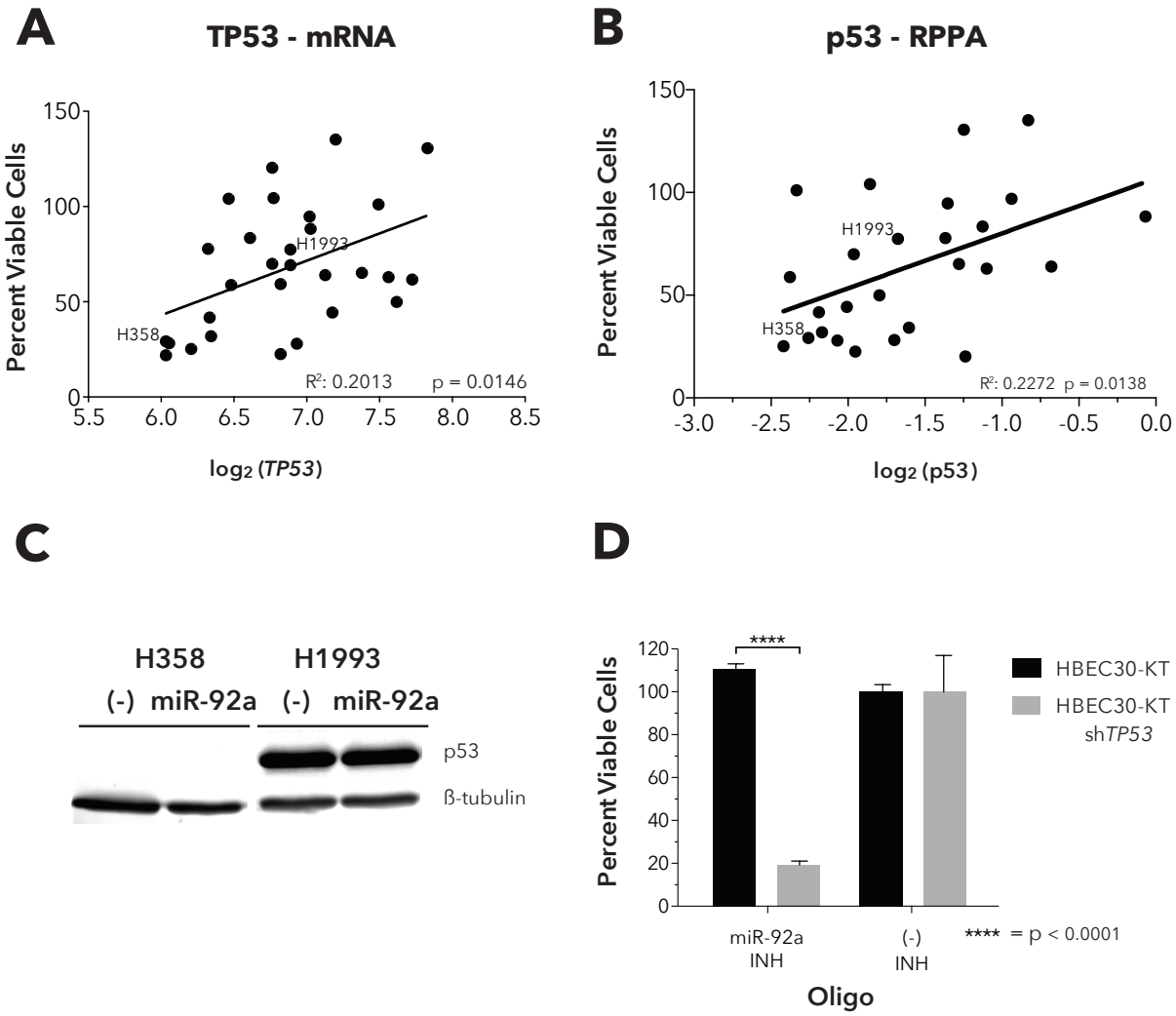


Figure 2.2.6 - Loss of p53 sensitizes lung epithelium-derived cells to miR-92a-induced toxicity. A) Scatter plot of *TP53* mRNA quantity versus cell viability after transient transfection of 50 nM of miR-92a. Presented as percent viable cells relative to a negative control LNA. Each point represents one cell line, with the Phase I cell lines labeled. Viability was measured 144h after miR-92a transfection. The black line of best fit was determined by linear regression. R^2 and p-value were determined by Pearson correlation test. **B)** Correlation of cell viability 144h after transfection with 50 nM miR-92a inhibitor with p53 protein abundance, as determined by Reverse Phase Protein Array (RPPA). Each point represents once cell line, with the Phase I cell lines labeled. R^2 and p-value determined by Pearson correlation test. Black line is the line of best fit as determined by linear regression. **C)** Immunoblot for p53 in H358 and H1993 cell lines 72h after a 50 nM transfection of either a negative control LNA ((-)) or the miR-92a inhibitor (92a). **D)** Viability of HBEC30KT and HBEC30KT-sh*TP53* 72h after transfection with 50 nM of either a negative control LNA ((-) inh.), the miR-92a inhibitor or the miR-1226* inhibitor. Error bars represent standard deviation, and p-values were determined by two tailed, unpaired Student's t-test.

2.2.4 Depletion of hsa-miR-92a is insufficient to induce toxicity in p53-depleted cell lines, but is sequence-specific.

One of the persistent caveats of any experimental approach involving short oligonucleotides is phenotypes caused by sequence-based, off-target effects [110]. One approach to determining the specificity of a given oligo involves using oligos with an alternative chemistry to see if the oligo reproduces the knockdown and phenotype observed. I employed double-hairpin RNA based miRNA inhibitors (Dharmacon miRidian miRNA inhibitors) to address this question due to their sequence composition and physical structure. I observed successful knockdown of the mature hsa-miR-92a locus by miR-92a inhibitors designed with either the LNA or RNA hairpin chemistries in both of the cell lines used in the first phase of the screen (Figure 2.2.7A-D). However, the RNA hairpin miR-92a inhibitor did not reproduce the toxicity of the LNA inhibitor in H358 cells (Figure 2.2.7E), and did not demonstrate any toxicity versus a control in H1993 cells (Figure 2.2.7F). As expected based on their correlated toxicities, the miR-1226* inhibitors replicated the phenotypes observed with the miR-92a inhibitors in both the sensitive and resistant cell line (Figures 2.2.7E, 2.2.7F).

To better establish the sequence dependence of the toxicity of the LNA-based miRNA inhibitors, I designed sequence-modified variants of the miR-92a and miR-1226* oligos, either scrambling the oligo while maintaining the nucleotide frequency or introducing evenly spaced mismatches across the length of the oligos (Table 2.2.2). I determined that the toxicity of the miR-92a & miR-1226* inhibitors was highly sequence specific.

After transient transfection with equal concentrations of either the toxic miR-92a or miR-1226* inhibitors, or one of the sequence modified variants, none of the sequence-modified oligos phenocopied the toxic oligos (Figure 2.2.7G). As expected, I did not observe a toxic phenotype with any oligo in H1993 cells (Figure 2.2.7F). This finding that toxicity of the miR-92a and miR-1226* oligos is specific to the initially designed sequence was recapitulated in the HBEC30KT model, with HBEC30KT-sh*TP53* sensitive only to the original miR-92a and miR-1226* oligos (Figure 2.2.7I) and HBEC30KT insensitive to the the toxic oligos and all of their derivatives (Figure 2.2.7J). I was motivated by the presence of two different inhibitors that recapitulated the same p53-attenuated toxicity in NSCLC, as well as the highly sequence-specific nature of the toxicity, to identify the target of the miRNA inhibitor relevant to the phenotype.

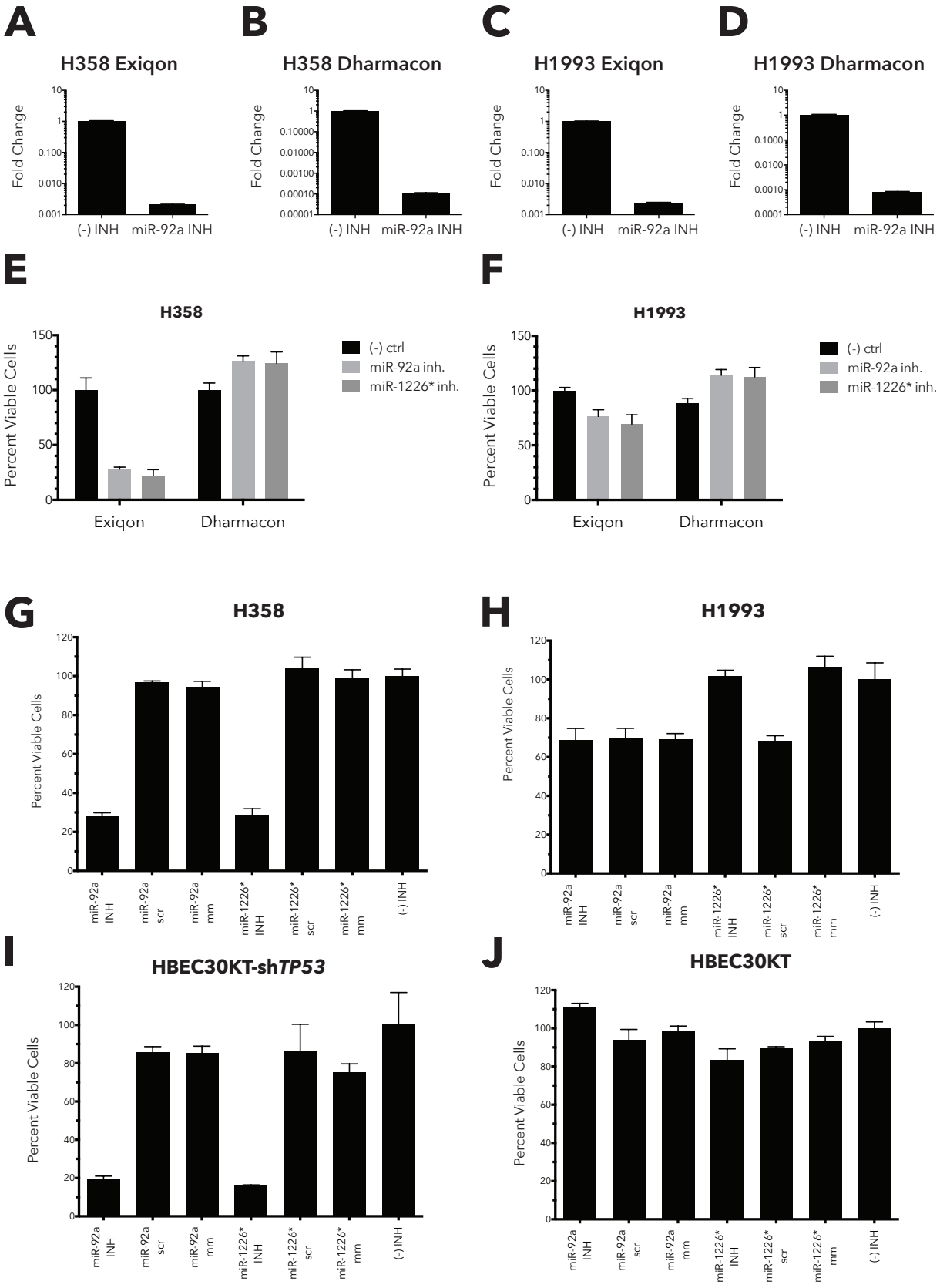

A
H358 Exiqon
Fold Change
(-) INH
miR-92a INH
B
H358 Dharmacon
C
H1993 Exiqon
D
H1993 Dharmacon
E
H358
Percent Viable Cells
(-) ctrl
miR-92a inh.
miR-1226* inh.
Exiqon
Dharmacon
F
H1993
G
H358
H
H1993
I
HBEC30KT-shTP53
J
HBEC30KT
miR-92a INH
miR-92a scr
miR-92a mm
miR-1226* INH
miR-1226* scr
miR-1226* mm
(-) INH

Figure 2.2.7 - Potency and toxicity of the miR-92a and miR-1226* inhibitors is sequence-specific. A) - D) qRT-PCR quantification of hsa-miR-92a in H358 and H1993 72h post transfection with either a negative control oligo ((-) INH) or a miR-92a inhibitor (92a INH). **A** and **C** were transfected with Exiqon LNA inhibitors, **B** and **D** were transfected with Dharmacon miRNA inhibitors. Bars represent the mean of 3 qRT-PCR replicates, error bars represent one standard deviation. Fold change is relative to negative control transfection in each panel. **E) and F)** Cell viability in H358 and H1993, respectively, 144h after transfection with 50 nM of negative control oligo, miR-92a inhibitor, or miR-1226* inhibitor. Labels on the x-axis identify the chemistry of the oligo pair, Exiqon or Dharmacon, assayed. Bars represent the mean of three replicates, error bars are s.d. Percent viable cells is relative to the negative control transfection in each case. **G - J)** Cell viability in H358, H1993, HBEC30KT-sh*TP53*, and HBEC30KT, respectively, 144h after transfection with either a control oligo ((-) INH), or a toxic or sequence-modified version of the miR-92a and miR-1226* inhibitors as described in **Table 2.2.2**. Bars represent the mean of 3 qRT-PCR replicates, error bars are s.d. Fold change is relative to negative control transfection in each panel.

Oligo Name	Sequence (5' -> 3')
miR-92a inhibitor	CAGGCCGGGACAAGTGCAAT
miR-92a scr.	GGCTAACAGCGCTAGAACGG
miR-92a mm	CAGCCCGCGACTAGTCCAAT
miR-1226* inhibitor	CCAGGCCTGCATGCCCTCA
miR-1226* scr	ACCACCGCGATCCGTGCTC
miR-1226* mm	CCACGCCAGCATGCCGTCA

Table 2.2.2 - Sequences of sequence-modified LNA oligonucleotides. Names and sequences of sequence-modified variants of the miR-92a and miR-1226* LNA oligos.

2.2.5 The miR-92a inhibitor can deplete the miR-17~92 primary transcript, as well as mature miRNAs from the miR-17~92 polycistron.

The miR-17~92 polycistron is a cluster of miRNAs with a shared origin on the same primary transcript (Figure 2.2.8A). This multi-miRNA cluster is dysregulated in many cancers -both amplification and over-expression are observed - in lung cancers, and ectopic expression of miR-17~92 increases proliferation in lung epithelium-derived cells [16,131]. Recent reports suggest a relationship between p53 and the miR-17~92 cluster; under hypoxic conditions p53 can transcriptionally repress miR-17~92 transcription [132]. Additionally, loss of the miR-17~92 cluster phenocopies *Dicer1* loss and is synthetic lethal with p53 and Rb in a mouse model of retinoblastoma, demonstrating a p53/miR-17~92 relationship in a tumor context [133].

In order to capture changes in the presence of the primary transcript of miR-17~92 I analyzed total RNA from miR-92a-sensitive and -resistant cells by qRT-PCR. Cells were transiently transfected for 72 hours with either control oligos or miR-92a inhibitors. Using the single stranded LNA inhibitors (Exiqon), I first observed a decrease in miR-17~92 primary transcript at two different positions on the transcript in H358 (Figure 2.2.8B). The Dharmacon RNA hairpin inhibitor to miR-92a, which can successfully knockdown mature hsa-miR-92a in the same cell line, showed no effect on the primary transcript (Figure 2.2.9C). Treatment with the miR-1226* inhibitor, which does not deplete mature miR-92a in H358 (Figure 2.2.9B), also led to depletion of the miR-17~92 primary transcript (Figure 2.2.9A). When I performed the same assay in a resistant cell line, H1993, I also observed a significant down-regulation of the miR-17~92 primary

transcript (Figure 2.2.8C). I observed the similar responses in both the sensitive HBEC30KT-sh*TP53* and the miR-92a inhibitor-resistant HBEC30KT, with significant decreases in miR-17~92 as detected at both locations on the primary transcript (Figures 2.2.8D, 2.2.8E). Consistent with previous reports demonstrating up-regulation of miR-17~92 in NSCLC, after comparing the relative abundance of the miR-17~92 primary transcript between tumor and normal cell lines I observe both significantly more miR-17~92 in tumor cells versus normal cells, and that sensitive cell lines had more abundant miR-17~92 than resistant cell lines (Figure 2.2.8F, 2.2.8G) [16]. These data demonstrate that the depletion of the miR-17~92 transcript is specific to the LNA-based inhibitor, consistent across sensitive and resistant cell lines, and can be reproduced by a miRNA inhibitor with a highly correlated toxicity to miR-92a.

I then assayed total RNA from four cell lines, H358, H1993, HBEC30KT and HBEC30KT-sh*TP53* for the abundance of the major mature miRNA loci from the miR-17~92 cluster (Figures 2.2.8H-K). When comparing mature miRNA abundance in cells transiently transfected with the miR-92a inhibitor versus the mismatched control I observed depletion of three mature miRNAs: hsa-miR-18a, hsa-miR-20 and hsa-miR-92a (Figure 2.2.8H-K). The miR-92a locus demonstrated the most complete knockdown in all cases. Mature miRNAs can have half-lives of hours to days in vivo; and variation in the cellular stability of other mature miRNA on the miR-17~92 locus may account for differences in relative abundance of mature miRNAs after transfection [134,135]. In the miR-92a resistant NSCLC cell line H1993 the only miRNAs observed to be depleted were hsa-miR-19a and hsa-miR-92a.

A similar pattern of mature miRNA depletion was observed in the HBEC30KT cell lines. Only hsa-miR-92a was depleted in the resistant HBEC30-KT (Figure 2.2.8K). All miRNAs on the miR-17~92 cluster were down regulated approximately two-fold in HBEC30KT-sh*TP53* (Figure 2.2.8J). The general trend of mature miRNA down regulation in sensitive cell lines is consistent across both tumor and and HBEC cell line model and consistent with the relationship between miR-92a toxicity and p53 abundance.

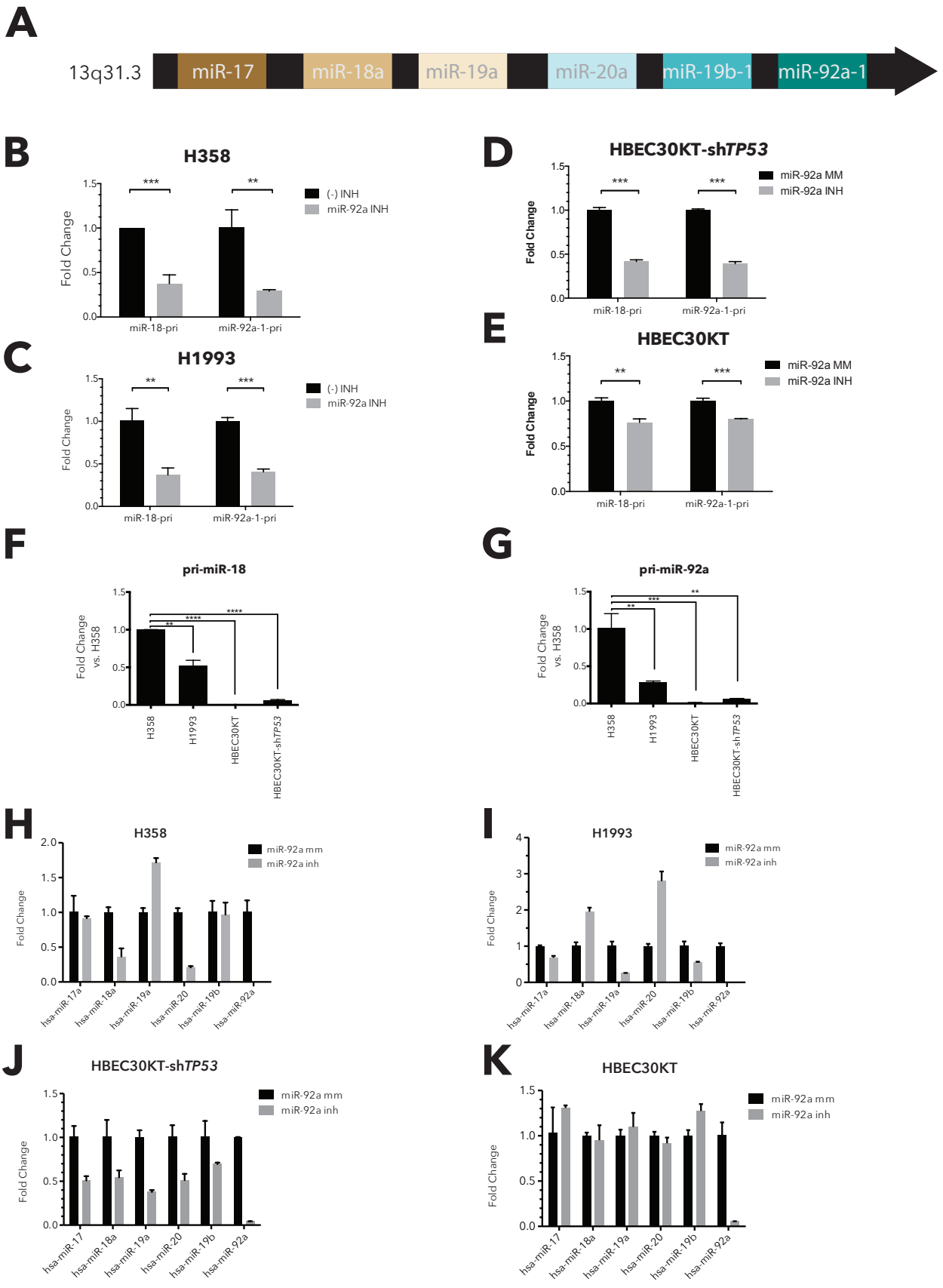

A
13q31.3
miR-17
miR-18a
miR-19a
miR-20a
miR-19b-1
miR-92a-1
B
H358
Fold Change
(-) INH
miR-92a INH
miR-18-pri
miR-92a-1-pri
C
H1993
D
HBEC30KT-shTP53
miR-92a MM
miR-92a INH
E
HBEC30KT
F
pri-miR-18
Fold Change vs. H358
H358
H1993
HBEC30KT
HBEC30KT-shTP53
G
pri-miR-92a
H
miR-92a mm
miR-92a inh
hsa-miR-17a
hsa-miR-18a
hsa-miR-19a
hsa-miR-20
hsa-miR-19b
hsa-miR-92a
I
J
hsa-miR-17
K

Figure 2.2.8 - The miR-92a inhibitor disturbs the abundance of both the primary transcript and individual miRNAs from the miR-17~92a polycistron. A) Illustration of the miR-17~92a polycistron on human chromosome 13q. **B)** qRT-PCR for the primary transcript (pri-miRNA) of the miR-17~92a polycistron in H358 at two locations: one in the miR-18a locus (18-pri) and one in the miR-92a-1 locus (92a-1-pri). RNA was collected 72h post transfection with either a negative control oligo or the miR-92a inhibitor. Bars represent the mean of 3 qRT-PCR replicates, error bars are s.d. Fold change is relative to negative control transfection in each panel. **C)** As in **B**, in H1993. **D & E)** qRT-PCR for the miR-17~92 primary transcript in HBEC30KT-sh*TP53* and HBEC30KT, respectively, after treatment with a 92a mismatch oligo, or the miR-92a inhibitor. **F)** Relative abundance of pri-miR-18 across cell lines. C_T values are normalized to loading control and then to H358. **G)** As in F, for the detection of pri-miR-92. **H-K)** qRT-PCR of mature miRNAs from the miR-17~92a polycistron 72h post transfection with either the mismatched Exiqon miR-92a oligo or the unmodified Exiqon miR-92a oligo in H358, H1993, HBEC30KT-sh*TP53* or HBEC30KT, respectively. Bars represent the mean of 3 qRT-PCR replicates, error bars are s.d. Fold change is relative to negative control transfection for each miRNA. p-values determined by Student's t test as described in methods, ** represents $p < 0.01$, *** represents $p<0.001$, **** represents $p<0.0001$.

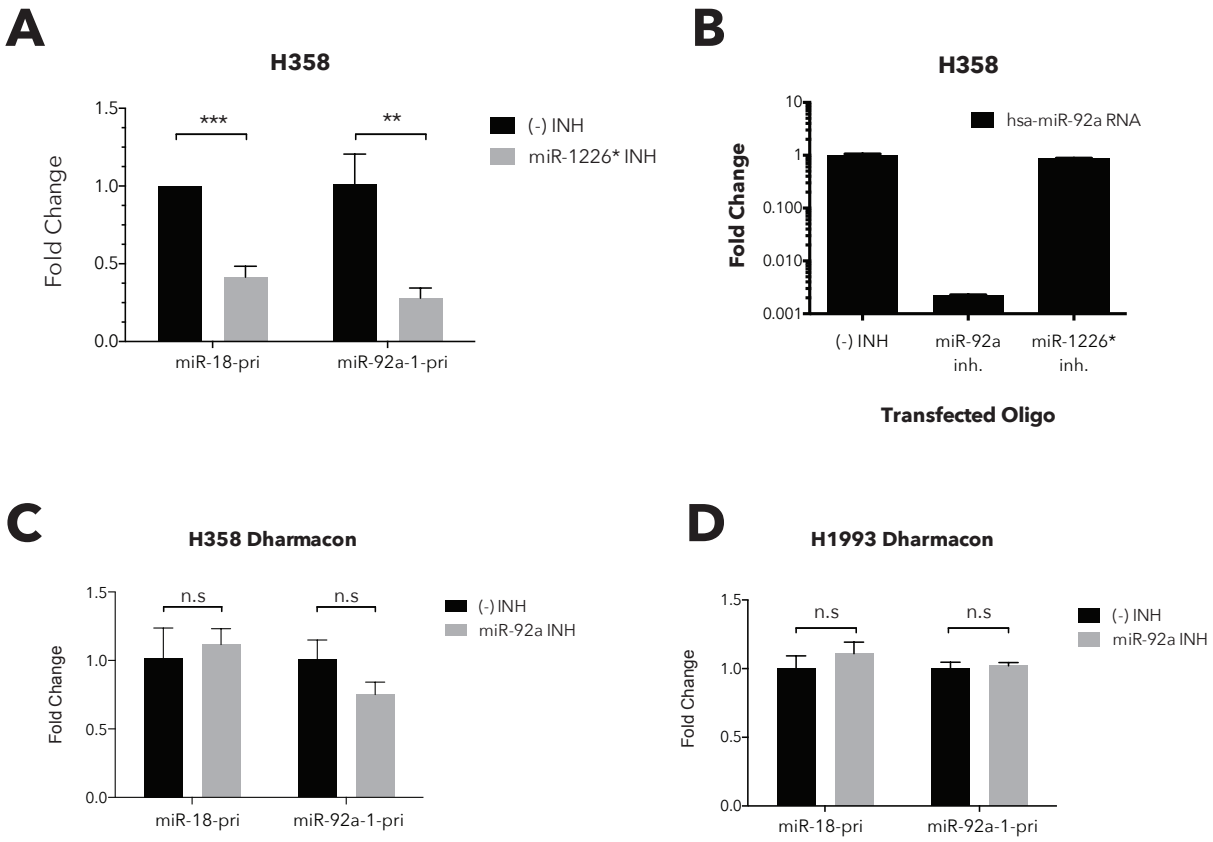


Figure 2.2.9 - Activity of miR-92a and miR-1226* inhibitors on the miR-17~92 primary transcript. A) pri-miR-18 and pri-miR-92 qRT-PCR on H358 mRNA 72 hours after miR-1226* inhibitor treatment. Bars are mean of 3 replicates, error bars represent s.d. p-values determined by Student's t test as described in methods, ** represents $p < 0.01$, *** represents $p<0.001$, **** represents $p<0.0001$. **B)** miR-92a qRT-PCR on H358 RNA 72 hours post transfection with either a control, miR-92a or miR-1226* inhibitor. Bars are mean of 3 replicates, error bars represent s.d. **C)** As in **A**, with the Dharmacon RNA inhibitor of miR-92a. **D)** As in **C**, save for treatment in the H1993 cell line.

2.2.6 The miR-92a inhibitor de-represses the miR-17~92 targetome in a context-dependent manner.

After observing the disruption of normal presence and processing of the miR-17~92 cluster and noting that high relative expression alone was insufficient for toxicity, I obtained transcriptional profiling data in a defined model of p53 loss in lung epithelium. The HBEC30KT model allowed us to investigate the effect of miR-92a transfection versus the mismatched control with a controlled difference in p53 abundance and function. I transiently transfected HBEC30KT and HBEC30KT-sh*TP53* with 50 nM of the miR-92a LNA inhibitor or its mismatched control for 48 hours and then collected RNA for microarray analysis.

In order to identify patterns of genetic reprogramming in cells after perturbation I first identified genes with a two-fold change in expression in the following conditions: mismatch-treated HBEC30KT-sh*TP53* versus mismatch treated HBEC30KT, miR-92a inhibitor-treated HBEC30KT versus mismatch-treated HBEC30KT, and miR-92a inhibitor-treated HBEC30KT-sh*TP53* versus mismatch-treated HBEC30KT-sh*TP53*. These conditions allowed for the identification of inherent differences in the cell lines after p53 knockdown (mismatch-treated HBEC30KT-sh*TP53* versus mismatch treated HBEC30KT) as well as differential responses of a resistant and sensitive cell line to the miR-92a inhibitor (miR-92a inhibitor-treated HBEC30KT versus mismatch-treated HBEC30KT, and miR-92a inhibitor-treated HBEC30KT-sh*TP53* versus mismatch-treated HBEC30KT-sh*TP53,* respectively). I observed that the magnitude of differential expression was higher after p53 loss than between the miR-92a inhibitor and mismatch

transfected conditions in either cell line (Figure 2.2.10A). The number of genes differentially expressed after p53 loss was 5,542, 12-fold higher than in HBEC30KT after miR-92a inhibitor treatment, and 6-fold higher than in HBEC30KT-sh*TP53* after miR-92a inhibitor treatment (Figure 2.2.10A). The relative frequency of up-regulation and down-regulation were similar in the two non-toxic conditions observed (mismatch-treated HBEC30KT-sh*TP53* versus mismatch treated HBEC30KT and miR-92a inhibitor-treated HBEC30KT versus mismatch-treated HBEC30KT) with 47.4% of differentially expressed genes up-regulated after p53 loss and 50.6% up-regulated in HBEC30KT after miR-92a inhibitor treatment (Figure 2.2.10B). Interestingly, in the miR-92a sensitive HBEC30KT-sh*TP53* there is an asymmetric response to the treatment, with 62.5% of genes up-regulated after treatment, a 11.9% shift from the observation in HBEC30KT (Figure 2.2.10B). When observing the behavior of genes up-regulated after miR-92a inhibitor treatment in HBEC30KT-sh*TP53* I observed that after p53 loss 75.5% of these transcripts either maintained their cellular abundance or decreased in abundance, suggesting that their maintenance or repression may be involved in survival of the cells after p53 knockdown (Figure 2.2.10B).

Based on the previous observation that the miR-92a inhibitor can deplete the miR-17~92 cluster I was interested in identifying whether or not I could observe a corresponding up-regulation of genes in the miR-17~92 target space. I therefore repeated the analysis of 2-fold up- and down-regulated genes using the subset of genes that are high confidence targets of the miR-17~92 cluster; these were determined using the 75th percentile of the TargetScan 6.2 context score as a cutoff [70]. The magnitude

of gene regulation observed after miR-92a treatment versus p53 knockdown was consistent with our observations in the superset of all genes tested on the microarray, with 530 genes differentially expressed after p53 knockdown, which was 20-fold higher than miR-92a inhibitor treated HBEC30KT and 8-fold higher than miR-92a inhibitor treated HBEC30KT-sh*TP53* (Figure 2.2.10C, Figure 2.2.10A). Additionally, the p53 knockdown and miR-92a treatment of HBEC30KT again showed similar frequencies of up- and down-regulated genes at 41.5% and 38.5%, respectively, of differentially expressed genes being up-regulated (Figure 2.2.10D). In contrast to this the miR-92a inhibitor-sensitive cell line, HBEC30KT-sh*TP53*, displayed an asymmetric up-regulation of differentially expressed genes with 71.4% of differentially expressed genes in the miR-17~92 targetome increasing in expression (Figure 2.2.10D). Additionally, 74% of the genes up-regulated after miR-92a inhibitor treatment in HBEC30KT-sh*TP53* were either maintained or down-regulated after p53 loss (Figure 2.2.10D). These observations demonstrate that asymmetric gene up-regulation after treatment or gene knock-down is an event specific to the the sensitive HBEC30KT-sh*TP53*, and this up-regulation is observed both in the set of all differentially expressed genes and in the high-confidence miR-17~92 targetome.

Having observed an up-regulation of gene expression after miR-92a transfection in HBEC30KT-sh*TP53* cells in both the superset of all genes and the subset of high confidence miR-17~92 target genes I attempted to identify patterns of gene regulation consistent with de-repression of the miR-17~92 targetome. The miR-17~92 cluster members hsa-miR-17, hsa-miR-19 and hsa-miR-92 were identified as significantly

inhibited upstream regulators in HBEC30KT-sh*TP53* after miR-92a inhibitor treatment by the "Upstream Regulators" analysis module of IPA (Ingenuity® Systems, www.ingenuity.com). (Table 2.2.4). Accounting for shared seed sequences of members of the miR-17~92 cluster (miR-17 and miR-20, miR-19a and miR-19b) the three miRNAs identified as upstream regulators account for every major mature miRNA on the cluster except for hsa-miR-18a. Additionally, the three miRNAs had negative activation z-scores which indicates that known downstream targets of an upstream regulator were expressed in a manner consistent with inhibition of the upstream regulator, supporting the previous observation of a de-repression of miR-17~92 targetome gene expression (Table 2.2.4, Figure 2.2.10D). When the up- and down-regulated targets of the miR-17~92 miRNA families are plotted with their relationship to the miRNAs, as well as with the expected and actual direction of regulation described, a mixed response is observed in HBEC30KT. Based on their Upstream Regulator activation z-scores the miR-17 and miR-19 families appear activated, while the miR-92a family appears inhibited (Figure 2.2.10E). In contrast, inhibition of the miRNA upstream regulators was consistent in HBEC30KT-sh*TP53* and of a greater magnitude than that observed in HBEC30KT (Figure 2.2.10F). The distinct effects on miR-17~92 target genes in the two cell lines suggests that the suite of target genes under active regulation by the miRNAs are distinct as well. The preference for gene up-regulation after depletion of this cluster suggests that its miRNAs may cooperatively play a role in promoting cellular survival after p53 loss in the bronchial epithelium.

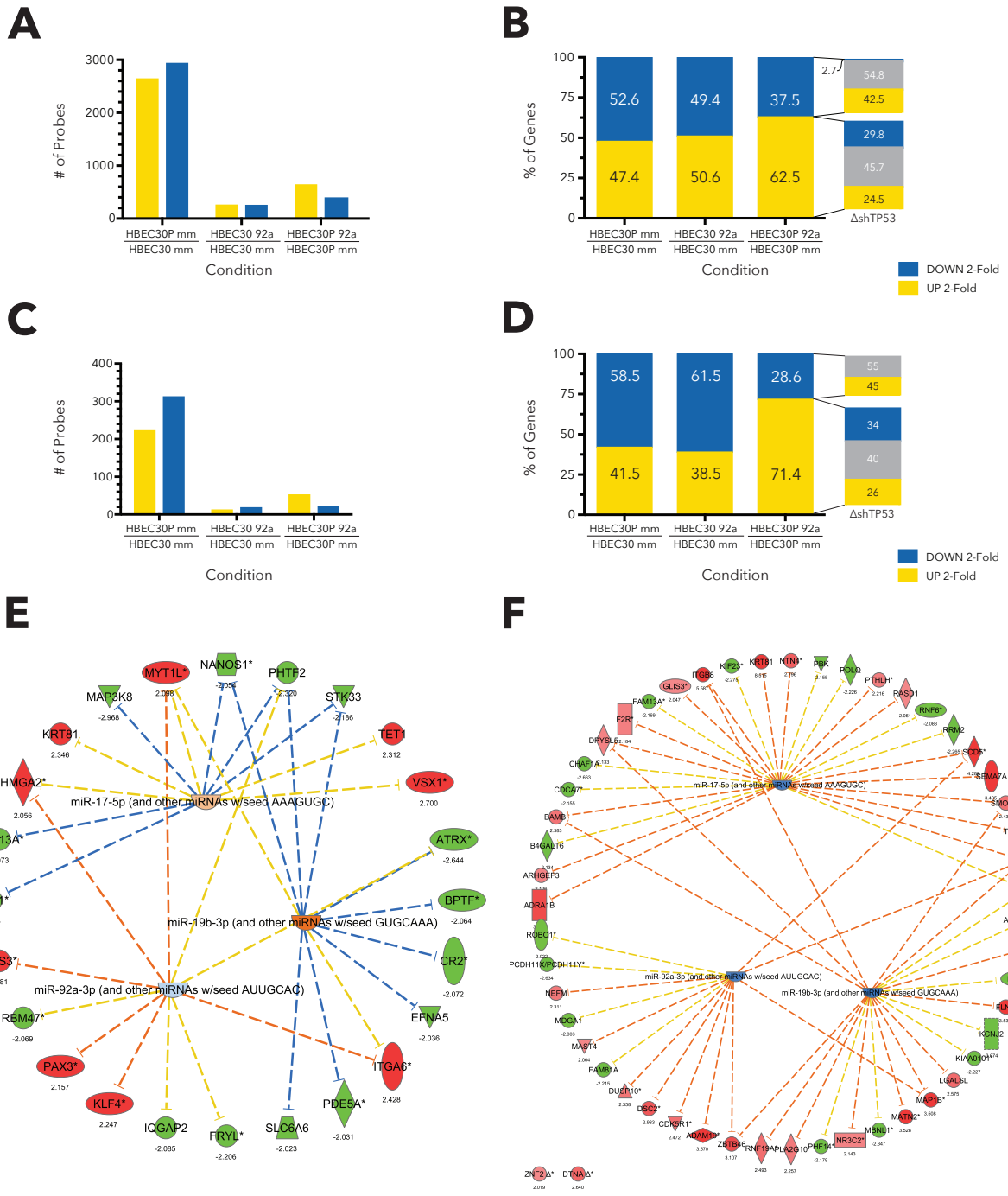

A
# of Probes
3000
2000
1000
0
HBEC30P mm / HBEC30 mm
HBEC30 92a / HBEC30 mm
HBEC30P 92a / HBEC30P mm
Condition
B
% of Genes
100
75
50
25
0
52.6
47.4
49.4
50.6
37.5
62.5
2.7
54.8
42.5
29.8
45.7
24.5
ΔshTP53
HBEC30P mm / HBEC30 mm
HBEC30 92a / HBEC30 mm
HBEC30P 92a / HBEC30P mm
Condition
DOWN 2-Fold
UP 2-Fold
C
# of Probes
400
300
200
100
0
HBEC30P mm / HBEC30 mm
HBEC30 92a / HBEC30 mm
HBEC30P 92a / HBEC30P mm
Condition
D
% of Genes
100
75
50
25
0
58.5
41.5
61.5
38.5
28.6
71.4
55
45
34
40
26
ΔshTP53
HBEC30P mm / HBEC30 mm
HBEC30 92a / HBEC30 mm
HBEC30P 92a / HBEC30P mm
Condition
DOWN 2-Fold
UP 2-Fold
E
MYT1L*
NANOS1*
PHTF2
MAP3K8
STK33
KRT81
TET1
HMGA2*
VSX1*
miR-17-5p (and other miRNAs w/seed AAAGUGC)
FAM13A*
ATRX*
ERC1*
BPTF*
miR-19b-3p (and other miRNAs w/seed GUGCAAA)
RGS3*
CR2*
miR-92a-3p (and other miRNAs w/seed AUUGCAC)
RBM47*
EFNA5
PAX3*
ITGA6*
KLF4*
PDE5A*
IQGAP2
FRYL*
SLC6A6
F
miR-17-5p (and other miRNAs w/seed AAAGUGC)
miR-92a-3p (and other miRNAs w/seed AUUGCAC)
miR-19b-3p (and other miRNAs w/seed GUGCAAA)

Figure 2.2.10 - Transcriptional profiling of HBEC30KT and HBEC30KT-sh*TP53* identifies a de-repression event driven by target genes of the miR-17~92 polycistron. A) Quantification of the number of genes increasing or decreasing greater than two-fold in the conditions tested. Samples transfected with the 92a mismatch oligo are labeled "mm," samples transfected with the miR-92a inhibitor are labeled "92a." HBEC30 = HBEC30KT, HBEC30p = HBEC30KT-shTP53, ratios are as described in the respective figures. **B)** Percentage representation of the values in **A**. Δsh*TP53* break-out represents the behavior of genes in each group in the “HBEC30p 92a/HBEC30p mm” column after knockdown of *TP53*. **C)** As in **A**, using the subset of genes identified by TargetScan as a target of a miRNA on the miR-17~92a cluster with a target score above the 75th percentile. **D)** Percentage representation of values in **C**, as in **B**. **E, F)** Relative expression of miR-17~92 target genes in HBEC30KT & HBEC30KT-sh*TP53*, respectively. Target genes identified as part of a miRNA regulatory network for miR-17/20, miR-19a/19b or miR-92a by Ingenuity Pathway Analysis. Fold change is relative to mismatch-treated control cells.

2.2.7 Inflammation-related pathways are up-regulated in HBEC30KT cells after p53 knockdown.

Part of the p53-mediated response to cellular stress involves transcriptional regulation of target genes. I observed that about 12% of the probes assayed by microarray responded greater than two-fold to the loss of p53 in HBEC30KT, without a preference for up or down regulation (Figure 2.2.10A). I then identified coherent genetic responses in this set of p53-reprogrammed genes. Via the "Upstream Regulators" analysis previously described an inflammation-like response was observed in the HBEC30KT-sh*TP53* cells. This observation comes from an analysis of the top 20 upstream regulators identified, where six of the regulators were identified as cytokines (TNF, IFNG, IL1B, OSM, IFNL1 and IFNA2), and all six were consistently identified as activated after p53 loss (Table 2.2.3). Two additional upstream regulators with known roles in inflammation and immune response, lipopolysaccharide and TLR3, were identified as activated upstream regulators (Table 2.2.3). The NFκB complex was also identified as an activated upstream regulator (Table 2.2.3). NFκB functions as a central hub of inflammation signaling, where many inflammation-related signaling pathways function through NFκB, and has a known antagonistic relationship with p53 [128,129].

Upstream Regulator	Molecule Type	Activation z-score	p-value of overlap
TNF	cytokine	7.142	1.11E-36
NKX2-3	transcription regulator	-5.01	1.20E-33
TGFB1	growth factor	-2.105	1.19E-32
IFNG	cytokine	8.079	1.28E-32
lipopolysaccharide	chemical drug	7.366	1.44E-32
beta-estradiol	chemical - endogenous mammalian	1.271	4.86E-25
ERBB2	kinase	1.86	1.49E-23
IL1B	cytokine	5.818	3.34E-22
TP53	transcription regulator	-3.853	5.76E-22
tretinoin	chemical - endogenous mammalian	0.696	7.04E-22
MAPK1	kinase	-4.712	1.38E-21
NFkB (complex)	complex	6.007	6.29E-21
TLR3	transmembrane receptor	5.119	1.22E-20
OSM	cytokine	4.083	1.80E-20
HRAS	enzyme	2.454	3.43E-20
IFNL1	cytokine	6.573	8.06E-20
dexamethasone	chemical drug	-6.854	1.50E-19
MAP3K7	kinase	1.672	5.39E-19
Cg	complex	1.14	6.68E-19
IFNA2	cytokine	6.286	1.74E-18

Table 2.2.3 - Twenty most significant upstream regulators in HBEC30KT - sh*TP53* versus HBEC30KT. As identified by Ingenuity Pathway Analysis of genes with two fold changes between conditions. Positive activation z-scores indicate predicted activation of the upstream regulator, negative z-scores predict inhibition of the upstream regulator.

2.2.8 The down-regulation of the miR-17~92 cluster mimics a cellular 1α, 25-dihydroxyvitamin D_3 response in the absence of p53.

The previously described consistent inhibition of mature miRNAs in the miR-17~92 cluster after p53 loss directed me towards identifying activated upstream regulators in the miR-17~92 targetome data set, as they would be activated at the same time the miR-17~92 target space was being up-regulated. The most significant upstream regulator identified in the miR-17~92 targetome in HBEC30KT-sh*TP53* after miR-92a inhibitor transfection was for 1α, 25-dihydroxyvitamin D_3 (Table 2.2.4). 1α, 25-dihydroxyvitamin D_3 is the biologically active form of Vitamin D_3, a fat-soluble secosteroid synthesized in the body, and is a key regulator of calcium and phosphorous regulation, as well as an immune modulator and regulator of cellular proliferation and differentiation [136-139]. 1α, 25-dihydroxyvitamin D_3 or its analogs have shown efficacy in reducing the proliferation of, and inducing apoptosis in, cancer cells [140-142]. Two members of the cytochrome P450 superfamily, *CYP27B1* and *CYP24A1*, are the rate-limiting regulators of 1α, 25-dihydroxyvitamin D_3 synthesis and degradation [142]. *CYP27B1* is the rate limiting enzyme in 1α, 25-dihydroxyvitamin D_3 synthesis, and *CYP24A1* is the rate-limiting enzyme in the catabolization of 1α, 25-dihydroxyvitamin D_3 after cellular import [142].

After establishing that in the set of genes consisting of the high-confidence miR-17~92 targetome, inhibition of the miR-17~92 cluster corresponded with a genetic response

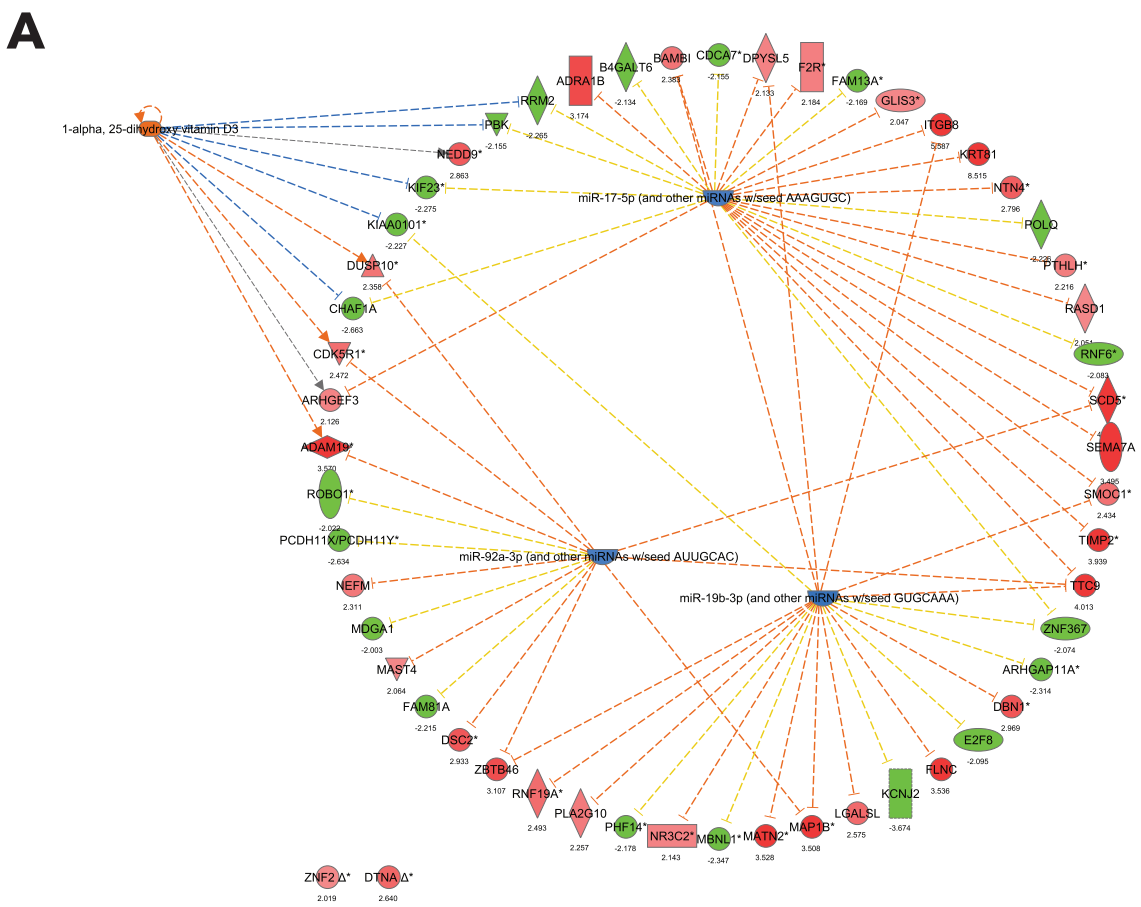
consistent with 1α, 25-dihydroxyvitamin D_3 activity I then plotted the hsa-miR-17, hsa-miR-19 and hsa-miR-92 target genes on the transcriptional network of 1α, 25-dihydroxyvitamin D_3 and observed an overlap between the miR-17~92 targetome and the 1α, 25-dihydroxyvitamin D_3 response network. I observed changes in gene expression consistent with activation of 1α, 25-dihydroxyvitamin D_3 signaling after miR-92a inhibitor transfection (Figure 2.2.11A). I also observed significant changes in the 1α, 25-dihydroxyvitamin D_3 metabolic and catabolic pathway components both between the HBEC30KT and HBEC30KT-sh*TP53* cell lines, and after miR-92a inhibitor treatment. Both *CYP27B1* and *CYP24A1* increased in the miR-92a inhibitor sensitive cell line (HBEC30KT-sh*TP53*) versus the miR-92a inhibitor insensitive cell line (HBEC30KT) (Figures 2.2.11B, 2.2.11C). *CYP24A1*, the enzyme that directly effects the intracellular half-life of 1α, 25-dihydroxyvitamin D_3, is decreased to basal levels post-miR-92a inhibitor treatment in HBEC30KT-sh*TP53* cells (Figure 2.2.11C).

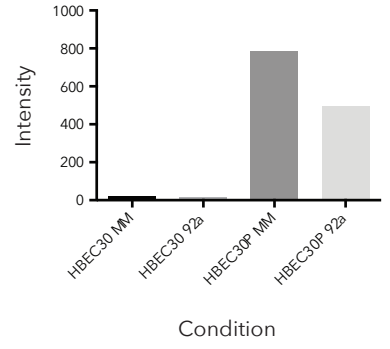
The concurrent increase in both the synthesis and degradation machinery for 1α, 25-dihydroxyvitamin D_3 in the HBEC30KT cell lines points to a symmetric adaptation after p53 loss to mitigate cellular 1α, 25-dihydroxyvitamin D_3 response. Increased 1α, 25-dihydroxyvitamin D_3 can induce *CYP24A1* transcription in a Vitamin D Receptor (VDR) mediated manner [142,143]. This is consistent with the up-regulation of both the Vitamin D_3 synthesis enzyme *CYP27B1* and the metabolic enzyme *CYP24A1* I observed in HBEC30KT-sh*TP53*.

Having employed the HBEC30KT pair of cell lines to identify the transcriptional changes observed in sensitive and resistant cells between two well-defined cell lines, I next wanted to ascertain the relevance of 1α, 25-dihydroxyvitamin D_3 in malignant, oncogenically transformed cells. I returned to our original miR-92a sensitive and resistant cell lines, H358 and H1993, and assayed their response to increasing concentrations of 1α, 25-dihydroxyvitamin D_3. I observed that the miR-92a sensitive cell line, H358, was significantly less viable after treatment with either 10 or 100 nM 1α, 25-dihydroxyvitamin D_3 than the miR-92a resistant cell line H1993 (Figure 2.2.12A). This result, in combination of our observation of a 1α, 25-dihydroxyvitamin D_3-like transcriptional response after miR-92a inhibitor treatment in sensitive cells (Figure 2.2.10F) demonstrates that the miR-92a inhibitor mimics the cellular response to 1α, 25-dihydroxyvitamin D_3.

Upstream Regulator	Molecule Type	Activation z-score	p-value of overlap
miR-17-5p (and other miRNAs w/seed AAAGUGC)	mature microRNA	-1.385	1.74E-15
miR-291a-3p (and other miRNAs w/seed AAGUGCU)	mature microRNA	-0.111	6.06E-11
miR-19b-3p (and other miRNAs w/seed GUGCAAA)	mature microRNA	-1.723	1.57E-10
1-alpha, 25-dihydroxy vitamin D3	chemical drug	2.791	4.53E-07
miR-92a-3p (and other miRNAs w/seed AUUGCAC)	mature microRNA	-1.367	8.39E-06

Table 2.2.4 - Five most significant upstream regulators in high-confidence miR-17~92 target genes HBEC30KT-sh*TP53*. As identified by Ingenuity Pathway Analysis of genes with two fold changes in expression after transfection of the miR-92a inhibitor versus the mismatched control. Target genes were filtered on the 75th percentile of the TargetScan context score. Positive activation z-scores indicate predicted activation of the upstream regulator, negative z-scores predict inhibition of the upstream regulator.


A
1-alpha, 25-dihydroxy vitamin D3
RRM2
PBK
NEDD9*
KIF23*
KIAA0101*
DUSP10*
CHAF1A
CDK5R1*
ARHGEF3
ADAM19*
ROBO1*
PCDH11X/PCDH11Y*
NEFM
MDGA1
MAST4
FAM81A
DSC2*
ZBTB46
RNF19A*
PLA2G10
PHF14*
NR3C2*
MBNL1*
MATN2*
MAP1B*
LGALSL
KCNJ2
PLNC
E2F8
DBN1*
ARHGAP11A*
ZNF367
TTC9
TIMP2*
SMOC1*
SEMA7A
SCD5*
RNF6*
RASD1
PTHLH*
POLQ
NTN4*
KRT81
ITGB8
GLIS3*
FAM13A*
F2R*
DPYSL5
CDCA7*
BAMBI
B4GALT6
ADRA1B
ZNF2Δ*
DTNAΔ*
miR-17-5p (and other miRNAs w/seed AAAGUGC)
miR-92a-3p (and other miRNAs w/seed AUUGCAC)
miR-19b-3p (and other miRNAs w/seed GUGCAAA)


B
D3 Synthesis
CYP27B1 Gene Expression
Intensity
1000
800
600
400
200
0
HBEC30 MM
HBEC30 92a
HBEC30P MM
HBEC30P 92a
Condition

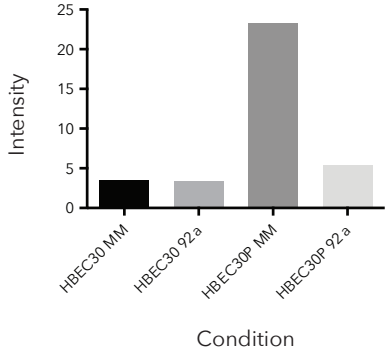
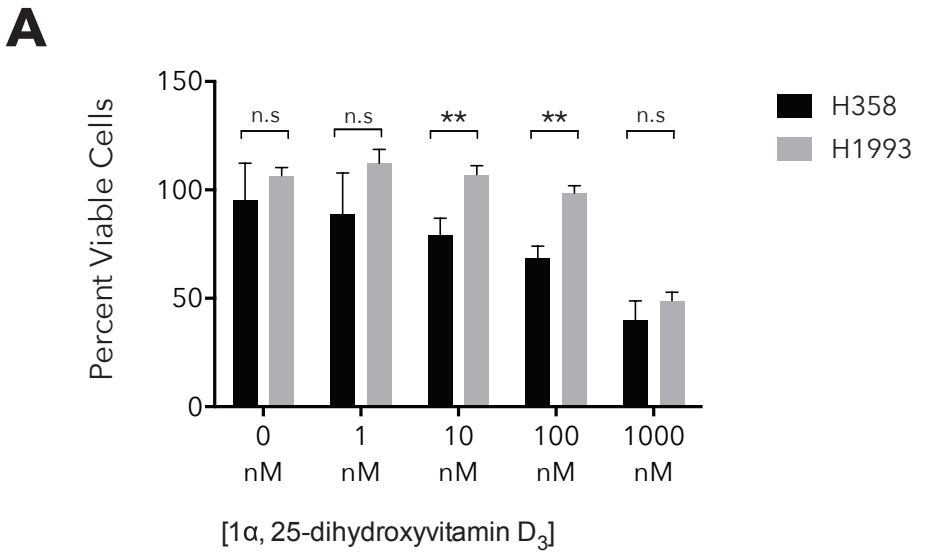

C
D3 Degradation
CYP24A1 Gene Expression
Intensity
25
20
15
10
5
0
HBEC30 MM
HBEC30 92a
HBEC30P MM
HBEC30P 92a
Condition

Figure 2.2.11 - 1α, 25-dihydroxyvitamin D_3 signaling is enhanced after depletion of the miR-17~92 cluster. A) Intersection of the high-confidence miR-17~92 targetome and 1α, 25-dihydroxyvitamin D_3 signaling pathway. Red colored values represent increased expression after miR-92a inhibitor treatment in HBEC30KT-sh*TP53*, green values represent decreased expression. Fold change values are listed underneath each item. Blue colored items are predicted to be inhibited upstream regulators, orange values are predicted to be activated. Orange and blue lines represent the same directionality for regulator:gene relationships, yellow lines represent observations incoherent with expected activity. **B)** Quantification of *CYP27B1* expression in all conditions tested. Bars represent mean intensity observed over three mRNA expression arrays. **C)** Quantification of *CYP24A1* expression in all conditions tested. Bars represent mean intensity observed over three mRNA expression arrays.



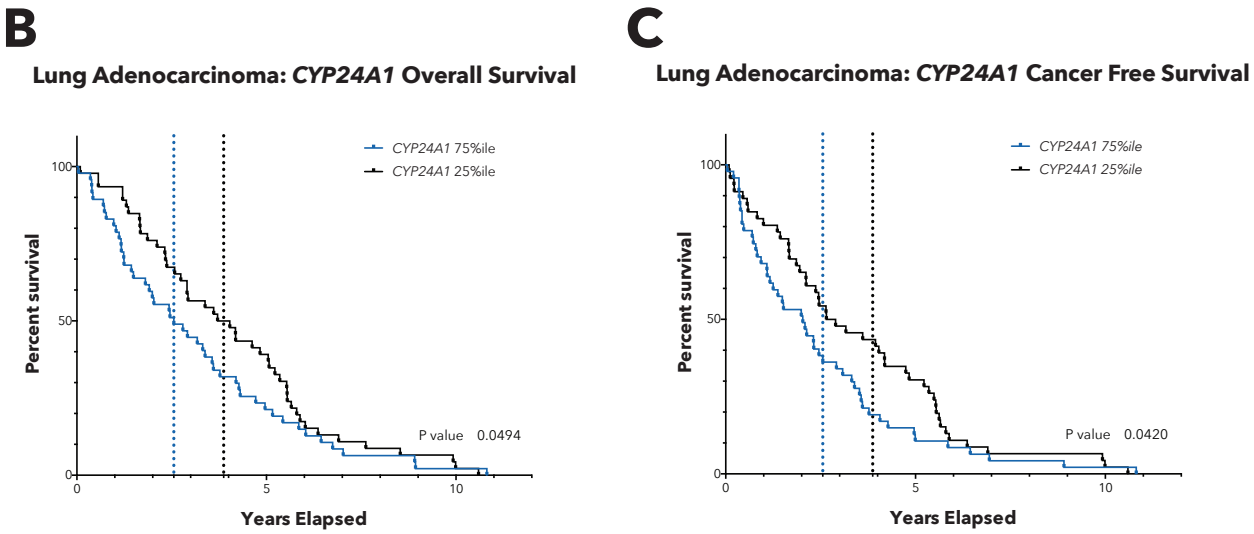


Figure 2.2.12 - 1α, 25-dihydroxyvitamin D_3 response in NSCLC. A) Cell viability of H358 and H1993 72h after 1α, 25-dihydroxyvitamin D_3 treatment. bars represent the mean viability across three replicates, bars represent s.d. ** = $p < 0.01$. **B)** Kaplan-Meier overall survival analysis for *CYP24A1* in lung adenocarcinoma patients. Blue line represents survival in samples in the upper quartile of *CYP24A1* expression, and the black line represents samples in the lower quartile of *CYP24A1* expression. Dotted vertical lines represent the median survival in the respective classes. **C)** Kaplan-Meier cancer-free survival analysis for *CYP24A1* in lung adenocarcinoma patients. Blue line represents survival in samples in the upper quartile of *CYP24A1* expression, and the black line represents samples in the lower quartile of *CYP24A1* expression. Dotted vertical lines represent the median survival in the respective classes.

2.2.9 *CYP24A1* is correlated with poor prognosis in lung adenocarcinomas

After establishing the connection between 1α, 25-dihydroxyvitamin D_3 response and miR-92a response in NSCLC cell lines I analyzed primary lung adenocarcinoma tumors for associations between 1α, 25-dihydroxyvitamin D_3 processing competency, as measured by *CYP24A1* abundance in tumors, and patient survival in lung adenocarcinomas (AC). Across 182 AC tumors I noted that high expressers of *CYP24A1*, as defined by expression levels above the 75th percentile, demonstrated significantly poorer overall and cancer-free survival than low expressers, defined as the 25th percentile and below (Figures 2.2.12B, 2.2.12C). Median overall survival was improved by 1.3 years in the low *CYP24A1* expressers (Figure 2.2.12B), and cancer-free survival improved by 0.863 years between the same groups (Figure 2.2.12C). These observations affirm previous observations that high *CYP24A1* correlates with poor survival in AC, and that 1α, 25-dihydroxyvitamin D_3 may have a preventative or palliative function in lung adenocarcinoma, and reaffirms our observation that the miR-92a inhibitor mimics the toxicity of a compound with a demonstrated beneficial effect in NSCLC [144,145].

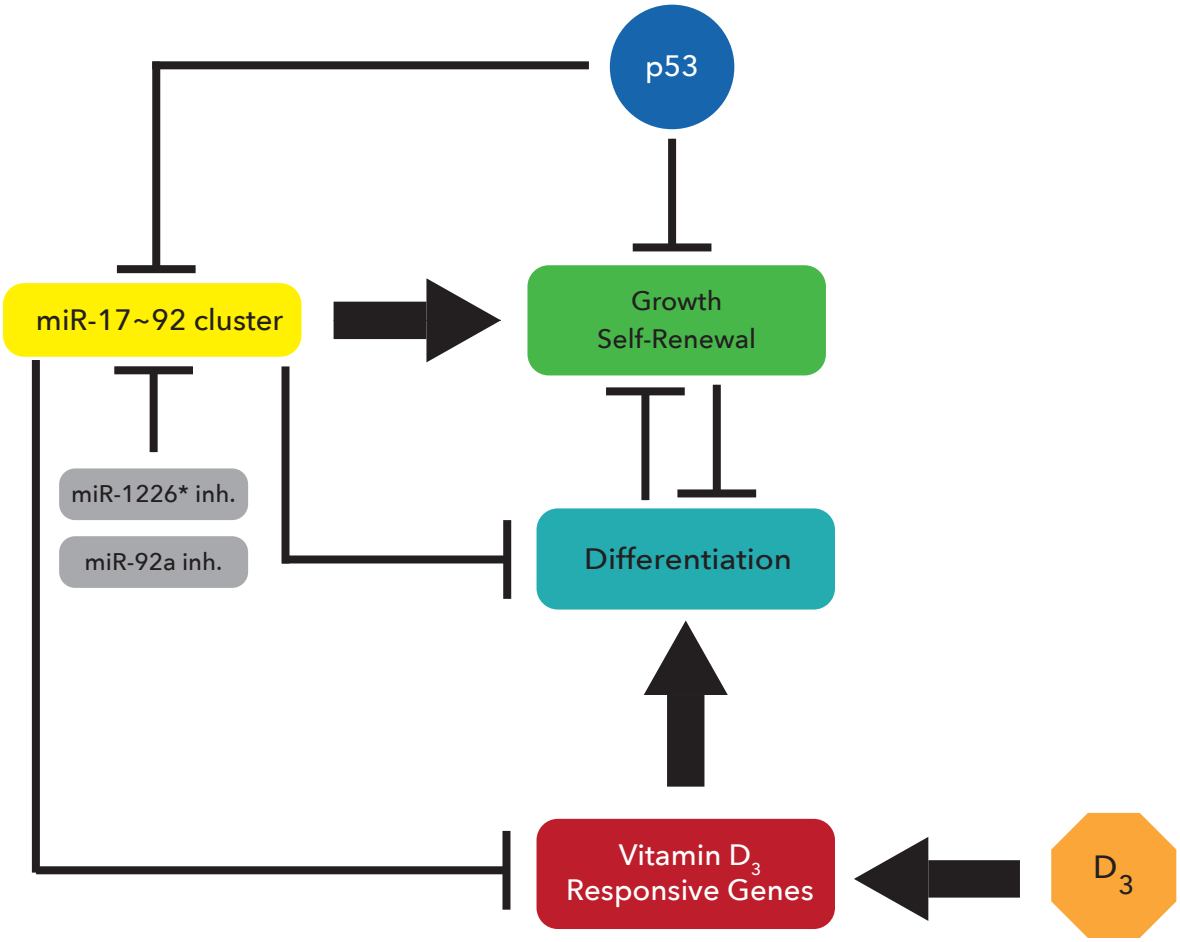


Figure 2.2.13 - Model for the interaction between p53, Inflammation, Vitamin D_3 and the miR-92a inhibitors. Working model of the miR-92a inhibitor, miR-17~92, p53 and 1α, 25-dihydroxyvitamin D_3 in NSCLC cellular growth and survival.

2.3 Discussion

Here I demonstrated that a miRNA inhibitor screen can identify inhibitors with selective toxicity across a panel of genetically diverse NSCLC cell lines. I then showed that the toxicity of the selectively toxic inhibitors anti-correlated with p53 abundance in these cell lines, that the knockdown of p53 in a resistant cell line was sufficient to sensitize it to the miR-92a inhibitor, and that the miRNA inhibitor for miR-92a acts in a sequence-specific manner to induce toxicity. I then observed that toxicity required activity beyond knockdown of the mature miR-92a product, and that this additional activity of down-regulating the miR-17~92 cluster resulted in a substantial up-regulation of the miR-17~92 targetome in a p53-depleted context-dependent manner. This up-regulation of the miR-17~92 targetome mimicked the 1α, 25-dihydroxyvitamin D_3 transcriptional response in a manner that was predicted by 1α, 25-dihydroxyvitamin D_3-induced toxicity (Figure 2.2.13).

By performing our primary screen in a pair of NSCLC cell lines with similar histological characteristics but with a distinct response to a specific treatment, *KRAS* depletion, I were able to identify miRNA inhibitors with several distinct patterns of toxicity across a panel of diverse NSCLC cell lines. In particular, I observed that the screen converged on one specific vulnerability, as two inhibitors I identified with selective toxicity in NSCLC cell lines but not in HBEC30KT, representative of normal lung bronchial epithelium. in fact displayed highly correlated toxicities. I were then able to identify p53 loss as a molecular mechanism that conferred sensitivity to the toxicity of the miR-92a and

miR-1226* inhibitors in NSCLC. The identification of a molecular mechanism also allowed us to employ a genetically-defined model system: the HBEC30KT cell line with wild-type p53 and with a constitutive knockdown of p53. I utilized this system to assess the changes in gene expression after p53 loss, to assess how those changes correspond to the availability of the miR-17~92 targetome, and identify the pathways and processes may be reprogrammed in response to p53 loss or miR-92a treatment in the sensitive and resistant contexts.

Surprisingly, the primary screen did not identify any miRNA inhibitors of this chemistry with high extreme potency in only one of the cell lines used in the first phase of the screen; the most toxic inhibitors were toxic in both cell lines. As previously discussed, lung adenocarcinomas tend to arise from perturbations in a limited set cellular processes and pathways, in contrast to a disease like serous ovarian carcinoma where few high-frequency mutations or copy number aberrations are observed [30,146]. The similar genetic reprogramming of the two lung adenocarcinoma cell lines used in the first phase of the screen may explain why I did not observe many inhibitors with dichotomous toxicity in our screen, as opposed to a miRNA mimic screen performed in serous ovarian carcinoma cell lines where many of the hits were highly idiosyncratic [Shields et al., 2013, in preparation].

The concept of specifically dysregulating a miRNA cluster with one compound in a manner that mimics a response to a different compound is an intriguing observation. On the one hand, it demonstrates the power of a miRNA-based genetic screen to identify

unexpected and novel vulnerabilities in cancer biology. This observation also demonstrates the importance of cellular context in miRNA activity: in the HBEC30KT model of p53 loss de-repression of the miR-17~92 targetome is only observed after the loss of p53.

In order for inhibition of one or several miRNAs to have a functional impact on the expression of the miRNAs' potential target genes two criteria need to be met. First, the gene(s) must be expressed and under miRNA regulation. Second, the inhibited miRNA(s) must be the key regulatory miRNA(s); there may exist redundant layers of miRNA activity on down-regulated genes that may not be overcome by the removal of a single miRNA. This is in contrast to miRNA gain-of-function studies where one miRNA can repress many miRNAs in a seed sequence-driven manner presumably without any influence from other cellular miRNAs.

miRNA:mRNA interaction sites are highly conserved across species, and the explanation offered for this is that miRNAs and their targets have evolved symmetrically in order to maintain homeostasis of regulatory circuits, suggesting that selective pressures may have led to the maintenance of the miR-17~92 polycistron as one single functional unit. Similar mechanisms have been observed for post-transcriptional miRNA regulation. mRNA transcripts that can act as post-transcriptional miRNA "sponges" were recently identified as systemic/network-based miRNA regulators [147]. Previous work has shown that the miR-17~92 cluster and its homolog cluster miR-106b~93 can act cooperatively to inhibit or activate multiple pathway members, such as p21 and *Bim* as

members of the TGFβ response pathway in gastric cancers [122]. This is consistent with our observation that multiple components of the miR-17~92 cluster cooperate to regulate a 1α, 25-dihydroxyvitamin D_3-like response in a p53-depleted context.

Specifically down-regulating a functionally interconnected miRNA regulatory network presents an opportunity to perturb a suite of related genes in order to achieve a desired outcome. However, the mechanism for the inhibitor:primary transcript interaction is yet to be determined, as the current knowledge of miRNA processing does not include nuclear export before Drosha processing, which would be necessary for canonical RISC-mediated knockdown. LNA-based antisense oligos have demonstrated potency both *in vitro* and *in vivo* in terms of effectively silencing longer mRNA transcripts, and our observation of miR-17~92 primary transcript depletion could be through an RISC-independent, antisense-like mechanism [148-150]. Conversely, the reports demonstrating that small RNAs and small RNA duplexes can modulate gene transcription by Agronaute-mediated interactions with RNA transcripts suggests an additional mechanism for small RNA mediated effects in the nucleus [151].

p53 aberrations are a common occurrence, present in up to 70-90% of tumors, depending on tumor type. Although the aberrations can take many forms including gains (and loss) of function, low p53 levels were clearly identified as a predictor of miR-92a inhibitor toxicity. The quantitative response of NSCLC cell lines to the miR-92a inhibitor in relation to p53 loss, coupled with the dramatic sensitization observed in HBEC30KT after p53 knockdown strongly suggests that the loss of p53 function is the event of

interest in our phenotype. As p53 mutation and loss are some of the most common genetic events in tumor biology, identifying a compound with antineoplastic effects in a low-p53 context represents a potential avenue towards p53-directed therapeutic interventions. Although there are currently no therapies targeted at loss of p53 regulation, and testing for a true loss-of-function of p53 activity in tumor cells may not be as straightforward as mutational profiling for other oncogenes or tumor suppressors that can only be activated or inactivated, the high frequency of p53 lesions in cancer makes it a strong candidate for therapeutic targeting.

While the benefits of 1α, 25-dihydroxyvitamin D_3 in lung cancer prevention and treatment are established, the ability of the miR-17~92 cluster to mimic 1α, 25-dihydroxyvitamin D_3 response in a p53-depleted context is a novel finding with therapeutic implications. The presence of hypercalcemia at therapeutic doses of 1α, 25-dihydroxyvitamin D_3 has limited its potential as an antineoplastic intervention.

Here I have described a mechanism for identifying miRNA-based interactions that can mimic the effect of a compound with known positive effects in cancer models. This approach has the potential to allow for the leveraging of the biological characterization of other small molecules with known therapeutic potential in model systems but poor clinical performance into miRNA and genetic therapeutic approaches that perturb the components of the drug's response necessary for a therapeutic effect without modulating the clinically deleterious nodes.

2.4 Methods and Materials

Cell Lines Cell lines beginning with “H” were established at the National Cancer Institute. Cell lines beginning with “HCC” and the immortalized human bronchial epithelial cells (HBECs) were established by the Hamon Center for Therapeutic Oncology Research at UT Southwestern Medical Center. All cancer cell lines, except for the HBECs and HCC4017, were grown in RPMI-1640 medium (Life Technologies, Rockville, MD) supplemented with 5% fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA). HBECs and HCC4017 were grown in ACL-4 medium supplemented with 2% serum [152,153]. All cell lines were grown in a humidified atmosphere with 5% CO_2 at 37˚C.

Reagents and Materials miRNA mimics siRNA oligos and miRidian miRNA inhibitors were obtained from Dharmacon (Chicago, IL). The miRCURY LNA™ microRNA Inhibitor Library - Human v12.0, was obtained from Exiqon (Denmark). Custom miRNA inhibitors for miR-92a, miR-1226* and their derivatives were also synthesized by Exiqon. The anti-rabbit antibody for p53 was acquired from Santa Cruz Biotechnology (Dallas, Texas) and the anti-mouse β-Tubulin antibody was acquired from Sigma Aldrich (St. Louis, MO). 1α, 25-dihydroxyvitamin D_3 was acquired from Sigma Aldrich (St. Louis, MO).

miRNA Inhibitor Screen & Analysis Cells were plated in 96-well format and reverse transfected with oligos for 72 hours and then supplemented with additional growth

media, followed by incubation for additional 72 hours. Cells, media, diluted Lipofectamine® RNAiMAX (Life Technologies) transfection reagent and diluted oligos were dispensed by MultiDrop Combi reagent dispenser (Thermo Scientific). Oligos were diluted to a final concentration of 50 nM, mixed and incubated according to the manufacturer's protocol with an appropriate quantity of diluted Lipofectamine® RNAiMAX (Life Technologies) transfection reagent before mixing with a single-cell suspension of the cell line of interest. Cell viability was determined using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega). Luminescence was quantified on a PerkinElmer EnVision plate reader (Waltham, MA). Raw luminescence values were then used to generate row-median normalized cell viability ratios using R [154]. The Bioconductor package "cellHTS2" [155] was further used to generate z-scores based on the normalized cell viability values.

Cell Viability Assay Cells were plated in 96-well format and, if necessary, transfected with oligos for 72 hours and then supplemented with additional growth media or growth media supplemented with 1α, 25-dihydroxyvitamin D_3, followed by incubation for additional 72 hours. Oligos were diluted for a final concentration of 50 nM, mixed and incubated according to the manufacturer's protocol with an appropriate quantity of diluted Lipofectamine® RNAiMAX (Life Technologies) transfection reagent before mixing with a single-cell suspension of the cell line of interest. Cell viability was determined using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega). Luminescence was quantified on a BMG LabTech PheraStar FS (Ortenberg, Germany) and analyzed in GraphPad Prism 6 (La Jolla, CA).

Quantitative Real-Time PCR Total RNA was prepared using the Ambion mirVana™ miRNA Isolation Kit (Life Technologies, Foster City, CA) 72 hours post transfection with 50 nM oligo. mRNA and miRNA levels were assessed by qRT-PCR using an ABI PRISM 7900 Sequence Detection System using predesigned TaqMan® primer and probe sets (Life Technologies, Foster City, CA). Reverse transcription and real-time PCR were performed according to the manufacturer's protocols and with equal quantities of total RNA input. RNU19 expression was used as a control for normalization of cDNA loading for miRNA assays, ACTB or GAPDH was used as a loading control for pri-miRNA and gene expression assays. Threshold cycle times (Ct) were obtained and relative gene expression was calculated using the comparative cycle time method.

Protein Analysis Samples were lysed in an SDS-TRIS lysis buffer. Protein concentration was determined using the Pierce BCA assay from Thermo Fisher (Rockford, IL), and equal quantities of cell lysate were electrophoretically resolved by SDS-PAGE. The protein was then transferred to PVDF membranes via the Bio-Rad Trans-Blot® Turbo™ Transfer System (Hercules, CA). Membranes were blocked with the LiCor Biosciences blocking buffer (Lincoln, NE) and probed with the previously described antibodies. Bound antibodies were detected with fluorophore-conjugated goat anti-mouse and goat anti-rabbit secondary antibodies (LiCor Biosciences), and visualized by an Odyssey® Infrared Scanner (LiCor Biosciences) and analyzed with the Image Studio software (LiCor Biosciences).

Acrylamide Gel DNA Electrophoresis Acrylamide gels were cast at 12% acrylamide in 1x Tris Borate EDTA (TBE) buffer using 10% (w/v) ammonium persulfate (APS). Gels were run in 1x TBE buffer at 30 mA for 70 minutes. Nucleic acids were stained by incubation with SYBR® Gold Nucleic Acid Gel Stain (Life Technologies, Foster City, CA) according to the manufacturer's protocol and bands were visualized using a 300 nm transilluminator.

Microarray Analysis HBEC30KT and HBEC30KT-*shTP53* cells were reverse transfected with 50 nM miR-92a mismatch or miR-92a inhibitor and plated in triplicate in 6-well plates. 48h post transfection total RNA was prepared using the Ambion mirVana™ miRNA Isolation Kit (Life Technologies, Foster City, CA). Total RNA was quality checked by Bio-Rad Experion (Bio-Rad, Hercules, CA). RNA was then processed and transcriptionally profiled on the Illumina HumanHT-12 v4 Expression BeadChip (Illumina, San Diego, CA) by the UT Southwestern Cancer Core (Dallas, TX) using standard methods. Briefly, The Ambion TotalPrep for Illumina (Life Technologies, Foster City, CA) kit was used to generate biotin-labeled cRNA, which was then denatured and hybridized to the arrays. After CY3-Streptavidin staining slides were scanned on an Illumina HiScan and signal intensities were summarized using GenomeStudio (Illumina). Background subtraction and quantile normalization were performed using the MBCB algorithm [156]. The post-normalization mean of the three replicates was used in all downstream analyses. miRNA target filtering was performed in R. Identification of upstream regulators in the expression data sets was performed using IPA (http://www.ingenuity.com).

Statistical Analysis Correlations were determined by Pearson correlation test, except for the miR-877* correlation which was determined by Spearman rank correlation. p-values determined by two-tailed, unpaired Student's t-test with $p < 0.05$ used as a standard cutoff for significance. Analyses were performed in GraphPad Prism (La Jolla, CA).

Reverse-Phase Protein Array (RPPA) RPPA was performed as previously described [103]**.** Briefly, equal amounts of filtered cell lysates were arrayed in triplicate on ONCYTE AVID™ nitrocellulose film slides (Grace Bio-Labs, Bend, Oregon), blocked and incubated with primary antibody at 4°C overnight. Slides were then washed and incubated with biotinylated secondary antibodies (Vector Laboratories, Burlingame, CA) for 30 min followed by blotting with Qdot 655–streptavidin conjugate (Life Technologies) for 30 min. Slides were scanned with a ProScanArray Microarray Scanner (PerkinElmer). Protein expression levels were quantified using MicroVigene™ (Vigene Tech, Carlisle, MA) software and normalized for differences in protein loading using Sypro Ruby™ (Life Technologies) protein stain signals obtained on a separate slide printed in the same batch.

Survival Analysis 182 lung adenocarcinoma tumor samples were collected, processed and summarized as described in [157]. Briefly, 10-20 μm thick serial sections of surgically resected NSCLC specimens were obtained using a Leica cryostat and homogenized using an Omni TH homogenizer (Omni International, Kennesaw, GA,

USA). Total RNA was isolated using TRIzol Reagent (Life Technologies). For microRNA profiling, the samples were labeled using the miRNA Complete Labeling and Hyb Kit, and hybridized to Agilent Human miRNA microarray version 3 chips (Agilent Technologies), which contains probes for 866 human and 89 human viral microRNAs based on miRBase v12.0 (http://microrna.sanger.ac.uk). miRNA expression levels were extracted using the Feature Extraction software (Agilent Technologies) and processed with the bioconductor package AgiMicroRna to correct for background, remove control and un-detectable sequences, normalize and summarize the data. Samples were then sorted by *CYP24A1* expression and the high and low quartiles were then compared by the Gehan-Breslow-Wilcoxon test in GraphPad Prism 6 (La Jolla, CA).

Chapter 3: Additional and Ongoing Projects

3.1 *KRAS* Dependence in NSCLC Cell Lines

3.1.1 Introduction

Constitutively activating *KRAS* mutations are one of the most prevalent and most potent oncogenic mutations in NSCLC [15]. Amplifications and/or activating mutations in codons 12, 13 or 61 occur in approximately 38% of lung adenocarcinomas, with a higher frequency in smokers versus never-smokers [15,30]. KRAS is a member of the Ras family of small GTPases oncogenes including HRAS and NRAS that, when activated, promote cell growth and survival primarily though the RAS/MEK/ERK and PI3K/AKT/mTOR signaling pathways [15,158]. Activating *KRAS* mutations prevent a transition to the inactive, GDP-bound state, and therefore activated KRAS continually sends pro-growth and pro-survival signals regardless of the presence of growth factors typically required for KRAS activation. This is especially evident in the case of EGFR tyrosine kinase inhibitors (TKIs), which inhibit the activity of constitutively active EGFR, an important growth factor receptor upstream of KRAS. NSCLC tumors harboring an activating *KRAS* mutation are almost entirely refractory to EGFR TKIs, presumably because EGFR-mediated activation of KRAS is no longer required once KRAS acquires an activating mutation [15].

While the high prevalence of *KRAS* mutations in NSCLC and even higher prevalences in other solid tumors such as pancreatic adenocarcinoma where nearly 90% of tumors contain activating *KRAS* mutations, KRAS has been a popular and rational molecule of

interest for the development of targeted therapies [159]. An early approach to inhibiting activated Ras proteins in cancer involved disrupting the post-translational modification, farnesylation, that allows Ras proteins to localize at the biologically active site at the cell membrane [159]. However, while some farnesyltransferase inhibitors showed activity against Ras proteins, and HRAS in particular, complications arose from the emergence of a secondary modification pathway, geranylgeranylation, that was sufficient to restore the biological localization and activity of Ras [159]. As there are many other proteins that require these modifications for normal cellular function, combined inhibition of farnesyltransferase and geranylgeranyltransferase activity was incredibly toxic [159]. More recent approaches to therapeutic targeting of Ras-driven tumors involve combined inhibition of parallel downstream pathways of Ras activity (PI3K & ERK inhibition), finding synthetic lethal interactions with *KRAS*-driven tumors, and identifying the relationship between KRAS activating mutations and tumor dependence on KRAS signaling [22,160-162].

Of particular interest to me was the 2009 report from Singh et al. where the authors first identified that tumor-derived cell lines harboring canonical *KRAS* mutations can respond variably to *KRAS* knockdown in cell culture [22]. They then stratified these *KRAS*-mutant cell lines into “KRAS Dependent” and “KRAS Independent” cell lines, and were then able to identify a gene expression signature that could, with a high degree of specificity, predict the KRAS-dependence of a cell line. The value added by this work is the understanding that not all *KRAS* mutations are equal, and that for the purposes of developing and deploying KRAS-targeted therapeutic approaches it may be imperative

to understand not only if a tumor has an oncogenic *KRAS* allele, but also whether that allele is relevant to that tumor's survival.

3.1.2 *KRAS*-mutant NSCLC Cell Lines Respond Paradoxically To EGFR TKI Inhibitors

After identifying *KRAS*-mutational status as the differentiating factor I would interrogate with the miRNA inhibitor library and screen described in Chapter 2, the question of which cell lines should be used in the primary screen arose. Certainly, one cell line should have a known activating *KRAS* mutation, and the other should not. However, I wanted to ensure that the cell line I chose as representative of *KRAS*-mutant NSCLC was in fact a good representative. Gefitinib is an EGFR TKI and is frequently used to treat *EGFR*-mutant NSCLC tumors, and *KRAS*-mutant NSCLC are known to be highly refractory to gefitinib treatment [15,163,164]. I chose two candidate cell lines based on *KRAS* mutational status and generated a gefitinib response curve for the *KRAS*-mutant H1155 and wild-type H1437.

Interestingly, the mutant cell line was slightly more sensitive to gefitinib than the wild-type cell line, with an observed EC_{50} of 10.22 μM in H1155 versus an EC_{50} of 49.90 μM in H1437 (Figure 3.1.1). Although neither cell line would be considered truly sensitive to gefitinib, this result led me to ask if I could identify better, and more representative cell lines of the *KRAS*-mutant and wild-type classes.

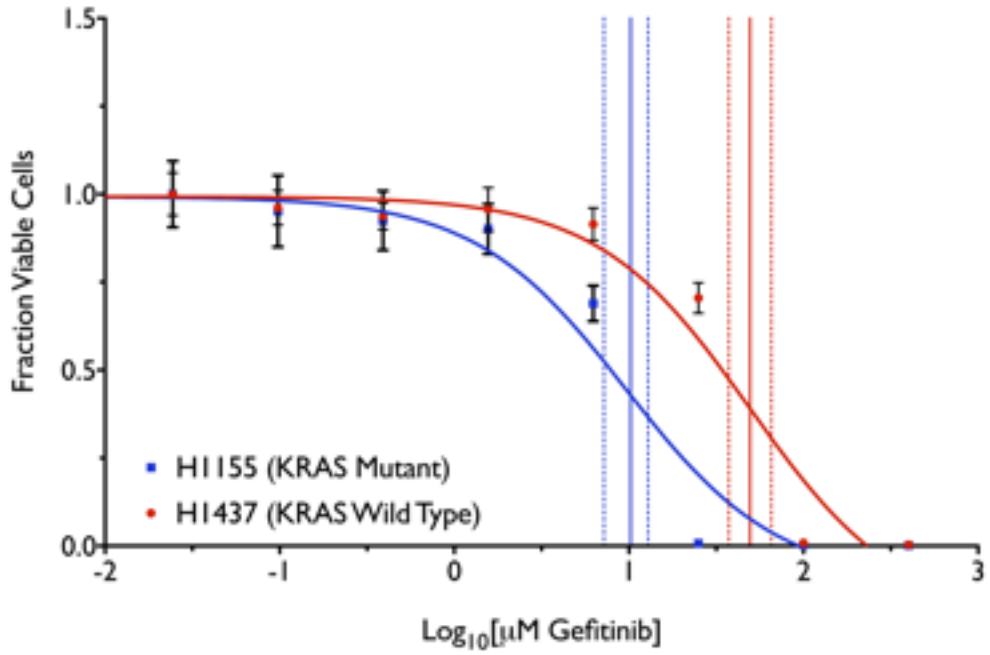


Figure 3.1.1 - Dose response curves after 72h of gefitinib treatment. A *KRAS*-mutant NSCLC cell line, H1155 and a wild-type cell line, H1437, were treated with gefitinib for 72 hours before viability was measured by luminescent ATP assay. Solid vertical lines represent the calculated EC_{50} for the color-matched cell line, and the dashed lines represent the calculated 95% confidence interval of the EC_{50}.

3.1.3 A Gene Expression Signature Identifies *KRAS*-dependent NSCLC Cell Lines

As previously described, a 2009 report from Singh et al. communicated that a gene expression signature could be used to classify *KRAS*-mutant cell lines as either dependent or independent of their *KRAS* mutation [22]. Dependent cell lines were defined as cell lines where cell viability was reduced by greater than 50% after RNAi-mediated knockdown of *KRAS* versus a control. Instead of individually testing all the available *KRAS*-mutant cell lines with *KRAS* RNAi, I applied the gene expression signature from the aforementioned report to first classify the mutant cell lines as dependent or independent.

Using the gene expression profiles of the NSCLC cell lines generated by John D. Minna lab I was able to classify *KRAS*-mutant NSCLC cell lines as *KRAS*-dependent or independent with only one misclassification, A549, based on the data reported in Singh et al. (Figure 3.1.2) [22]. Interestingly, when I included immortalized normal lung epithelium cells they were strongly clustered with the *KRAS*-dependent cell lines, which is consistent with the observation that *KRAS*-dependent cell lines appeared more epithelial than the *KRAS*-independent cell lines (Figure 3.1.2) [22].

In order to experimentally verify my findings I chose two cell lines one *KRAS*-dependent and one *KRAS*-independent cell line for treatment with an siRNA pool to *KRAS* and assess the impact of *KRAS* knockdown on survival (Figure 3.1.3). I saw that the predicted *KRAS*-dependent cell line, H358, responded as expected to *KRAS*

knockdown with a 50% decrease in viability, whereas the *KRAS*-mutant-but-independent cell line, H157, did not respond as robustly to *KRAS* knockdown (Figure 3.1.3).

As the goal of this study was to identify a suitable, representative *KRAS*-driven NSCLC cell line and a *KRAS*-agnostic second NSCLC cell line, I then chose a *KRAS*-wild type cell line, H1993, and assessed its response to *KRAS* depletion in comparison to the mutant and dependent cell line H358. The aim was to find a pair of cell lines where only the *KRAS*-mutant and dependent cell line was the only one that responded to *KRAS* depletion. I observed that, again, only H358 responded to *KRAS* knockdown, and I was also observed significant and similar knockdown of *KRAS* mRNA in both cell lines (Figure 3.1.4). Therefore, I proceeded to begin the screen in Chapter 2 with H358 and H1993 as the two genetically distinct but histologically related cell lines.

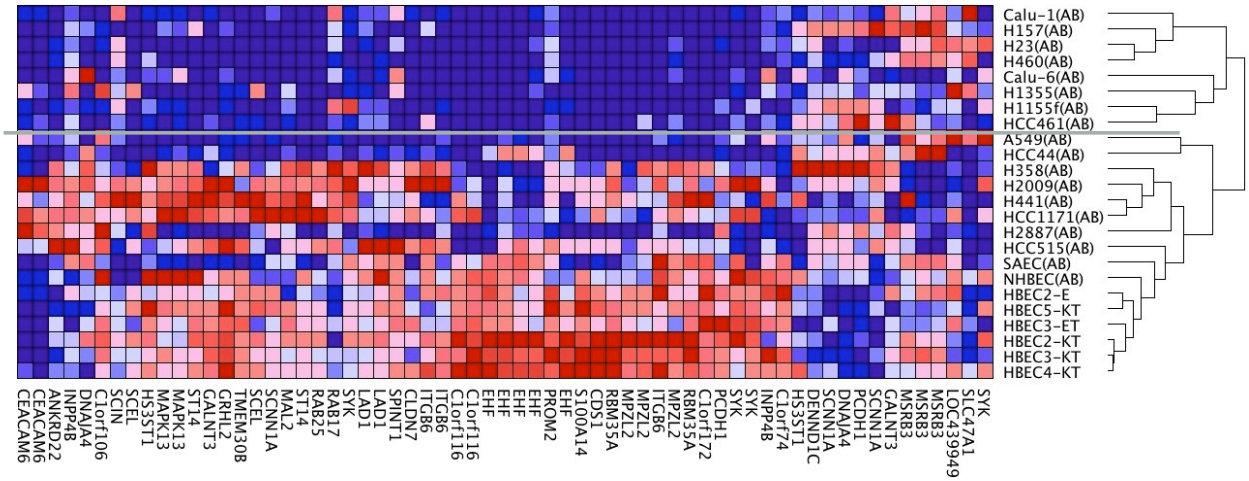


Figure 3.1.2 - Gene expression signature of *KRAS* dependence in NSCLC cell lines. Heat map of genes in the *KRAS* dependence gene expression signature. Color scale is, from lowest to highest, deep blue to dark red. Dividing line is the boarder between the *KRAS* independent cluster, top, and the *KRAS* dependent cluster, bottom. Cell line names label each row, gene names for each probe label each column.

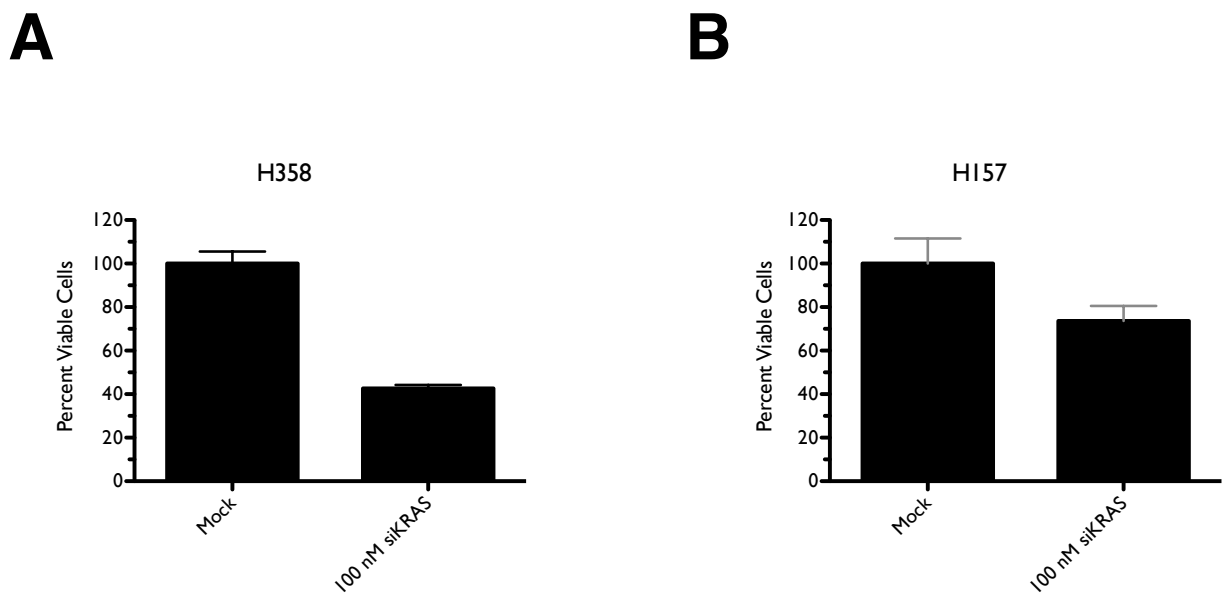


Figure 3.1.3 - Knockdown of *KRAS* in an *KRAS*-dependent and *KRAS*-independent cell line. A) Cell viability of H358 six days after *KRAS* knockdown or mock transfection control. Bars are the mean of 3 replicates, error bars represent s.d. **B)** As in **A**, in the *KRAS*-independent H157 cell line.

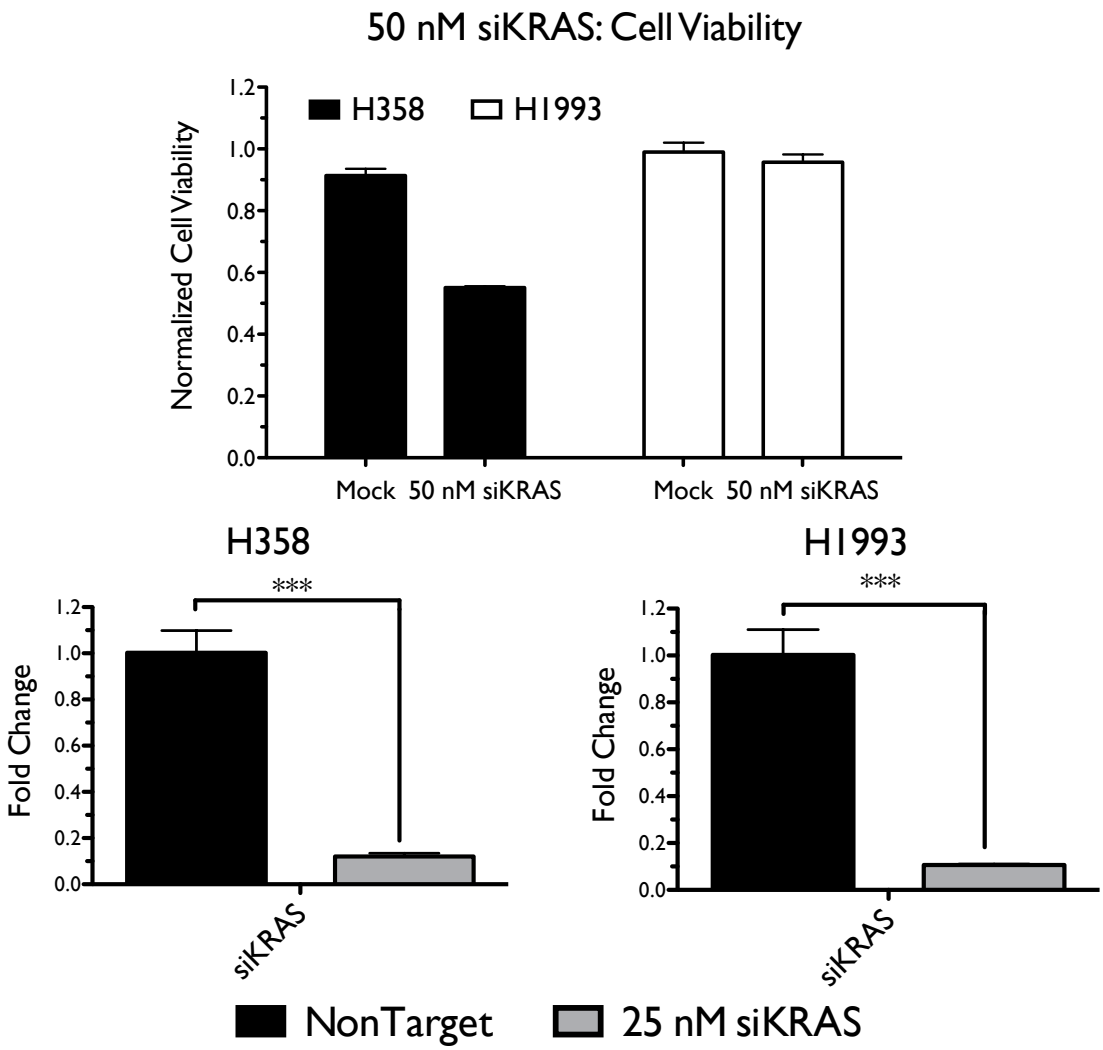


Figure 3.1.4 - Response to *KRAS* depletion in the two primary screen cell lines. Top, Cell viability after mock or si*KRAS* treatment for 6 days. Black bars represent results in H358, clear bars represent H1993 results. Bottom, qRT-PCR for *KRAS* 72 hours after treatment with a negative control oligo pool or si*KRAS* oligos. Fold change in H358, bottom left, and H1993, bottom right, relative to control treated. *** = $p < 0.001$, determined by Student's t-test.

3.1.4 Conclusions

After identifying the variable response of *KRAS*-mutant NSCLC cell lines to perturbations of KRAS signaling pathways, I was able to use an existing expression signature for *KRAS* dependence to confirm and pre-classify NSCLC cell lines as either *KRAS* dependent or independent. As *KRAS* biology is of great interest and importance in cancer biology, it will be important for ongoing research into the activity of mutant *KRAS* to identify the necessity of *KRAS* in the tumors or cell lines of interest, as associations observed in situations where *KRAS* is not required may be a result of other biologies.

Unfortunately, the results of the miRNA inhibitor screen did not yield the insights into miRNA dependency in *KRAS*-driven NSCLC that I anticipated. After I identified miRNA inhibitors with selective toxicity in the NSCLC cell line panel, none of the inhibitors had toxicity patterns that mapped coherently to either *KRAS* mutational status or dependence. This result is likely a consequence of a study underpowered for this particular phenotype with only one data point in each class. It is also possible that miRNA loss of function may not be a critical event in the maintenance of KRAS-driven oncogenesis in NSCLC, although previous work suggests that the requirement of miRNAs in KRAS-driven NSCLC is very miRNA-specific [75].

3.1.5 Methods and Materials

Cell Culture & Transfections Tumor cell lines were grown in RPMI-1640 medium (Life Technologies, Foster City, CA) supplemented with 5% fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA). HBECs and SAECs were grown in KSFM medium (Life Technologies, Foster City, CA). All cell lines were grown in a humidified atmosphere with 5% CO2 at 37˚C. Transfections were performed by as described in Chapter 2. Dharmacon siKRAS OnTarget PLUS siRNA pool (Thermo Scientific, Pittsburgh, PA) was used for knockdown of *KRAS* at concentrations described in figures.

Hierarchical Clustering Analysis Affymetrix U133 genomic profiles of NSCLC and HBEC cell lines were graciously supplied by Dr. John D. Minna. The gene expression signature for *KRAS* dependence was derived from the the previously mentioned study and translated to the arrays used by JDM [22]. Hierarchical clustering was performed using the “Hierarchical clustering” GenePattern module using the default settings, with column normalization [165]. Heat map images were then exported using the “Hierarchical cluster viewer” module.

Quantitative and Real-Time PCR - qRT-PCR assays were performed as described in Chapter 2.

Gefitinib Dose Responses Cells were plated in 96-well plates at an appropriate density to avoid confluence at the endpoint. After 72 hours Gefitinib (LC Labs, Woburn, MA) was applied to cells in increasing concentrations to cells. Cells were allowed to

grow for 72 hours before cell viability was quantified by CellTiter-Glo® Luminescent Cell Viability Assay (Promega). Luminescence was quantified on a BMG LabTech PheraStar FS (Ortenberg, Germany).

3.2 Big Data and Data Visualization In Modern Cancer Research

3.2.1 Introduction

The scalability and power of newly developed methods for genetic analysis have radically transformed the scope of investigation in cancer research [166,167]. In a dramatic example of the impact of high-density, high-resolution genetic data on turning identified oncogenic driver mutations into targeted therapies is the development time of the EML4-ALK inhibitor crizotinib. One of the earliest examples of an approved targeted therapy, the BCR-ABL inhibitor Gleevec, took 41 years from identification of the Philadelphia chromosome to the launch of the inhibitor [166]. Crizotinib, on the other hand, went from repositioning an existing compound to Phase II trials in only three years [166]. While variation, methylation, copy number and expression profiling of tumors on the scale necessary for reasonable sample sizes for rarer events is technically feasible, several practical considerations prevent many researchers from generating their own comprehensive independent data sets: well annotated primary tumor tissues are a scarce and high-value resource, and data generation and analysis requires considerable investment in each technological platform.

The preferred solution to the aforementioned issued appears to have emerged in the form of national & international consortia modeled after the Human Genome Project where a collection of researchers all contribute their expertise & resources to generate a shared & ideally freely accessible resource. A particularly relevant example in the cancer genetics context is The Cancer Genome Atlas (TCGA). TCGA has collected genome-wide mutation, mRNA, miRNA, copy number and methylation profiles across

greater than twenty tumor types, with hundreds of tumors and matched or unmatched normal samples for each tumor type, and have made all non-identifying information freely available to all. However, while the ready availability of such stores of data can enable many insightful inquiries, new challenges arise in terms of collecting data, creating focused queries and conveying the results in an interpretable, meaningful manner.

3.2.2 Circos, A Tool For Visualizing High-Density Genomic Data

Continuing improvements in the speed, quality, and accessibility of whole-genome analysis has brought with it the challenge of displaying the generated data in an informative and accessible manner. Existing conventions for conveying the results of experiments focused on a small, tractable number of individual observations do not scale to genome-wide resolution. Additionally, data visualizations based off of a standard, rectangular two-axis system do not allow for the display of spatial and composition relationships either within or outside of the dataset.

Recently, a new tool, Circos, has appeared to simplify the process of both conveying the salient features of genome-scale, multi-dimensional data sets [168]. Built in the readily accessible Perl language and based off of highly customizable & readable configuration files, Circos allows for the layering of multiple genome-wide data sets in a manner that can, for example, display regions of the genome where copy number aberrations correspond with changes in gene expression, or where genetic variation corresponds with expression changes. In addition to the programatic flexibility provided by Circos in

the types of plots that can be generated, and the programatic rules that can be applied to the display of data on those plots, the use of a circular, chromosome-based ideogram as the axis of reference [168]. Various plot types, including scatter plots, line graphs, histograms and tiles are then layered within the context of the circular, genomic axis, which additionally allows for the use of arc-like “ribbons” to display relationships between regions on the circular axis. Here I will display the results of a collaboration where this methodology was employed to bring added clarity to the analysis of genome-wide data analysis.

This example is the result of a collaboration with Benjamin Shields, who wanted to identify genetic differences between two serous ovarian carcinoma cell lines, PEO1 and PEO4, derived longitudinally from the same patient. After collecting RNA-Seq, single nucleotide variation and copy number variation data on the two cell lines there were two challenges to address: viewing the data in a manner that added insight to the differences between the cell lines, and conveying the significance and magnitude of the collected data in a visually tractable manner (Figure 3.2.1).

After several rounds of iteration a decision was made to move the karyotype (genomic position) from the outermost track to the middle of the circle, where it would still convey location, as well as serve as a divider between data types. The “PEO1/PEO4 mRNA” track had the highest data density, and was therefore placed on the largest radius track. Data on this track were $\log_2$ normalized RPKM+1 ratios. Cell-line specific color was used to both identify the cell line with higher expression via a colored background, as

well colored bars to identify a greater than two-fold change between cell lines. A quantitative data track was not used for single nucleotide variants, as this is a binary presence/absence data call. Therefore, colored tick marks were used to identify the presence of an SNV, and a neutral black was used to identify SNVs in common, and assumed to be inherited from the tumor's cells of origin.

Copy number variations were perhaps the most difficult of the data sets to plot in a satisfying manner. Early attempts to display these data tracks utilized a color scale to convey the magnitude of copy gain or loss in each cell line; however, this made distinguishing between gains and losses difficult without either losing the losses in to the white space or breaking the cell-line specific color scheme. Instead, the color scale for magnitude was eliminated and gains and losses were separated into independent tracks. While sacrificing the dynamic range of the copy number gains, the separation of gains and losses allows for a clear visualization of shared and independent regions of copy gain and loss across the genomes of the two cell lines. For example, the amplification of one arm of chromosome 8 in both cell lines is clear in this configuration, and a corresponding over-expression in PEO1 of the amplified region of the X chromosome in PEO1 was observed.

Here, the genome-wide visualization allowed for several insights. Copy number losses were far more prevalent in both cell lines than copy number gains and were widely distributed across the genomes of both cell lines. The configuration file used for generating this figure is available as Appendix B.

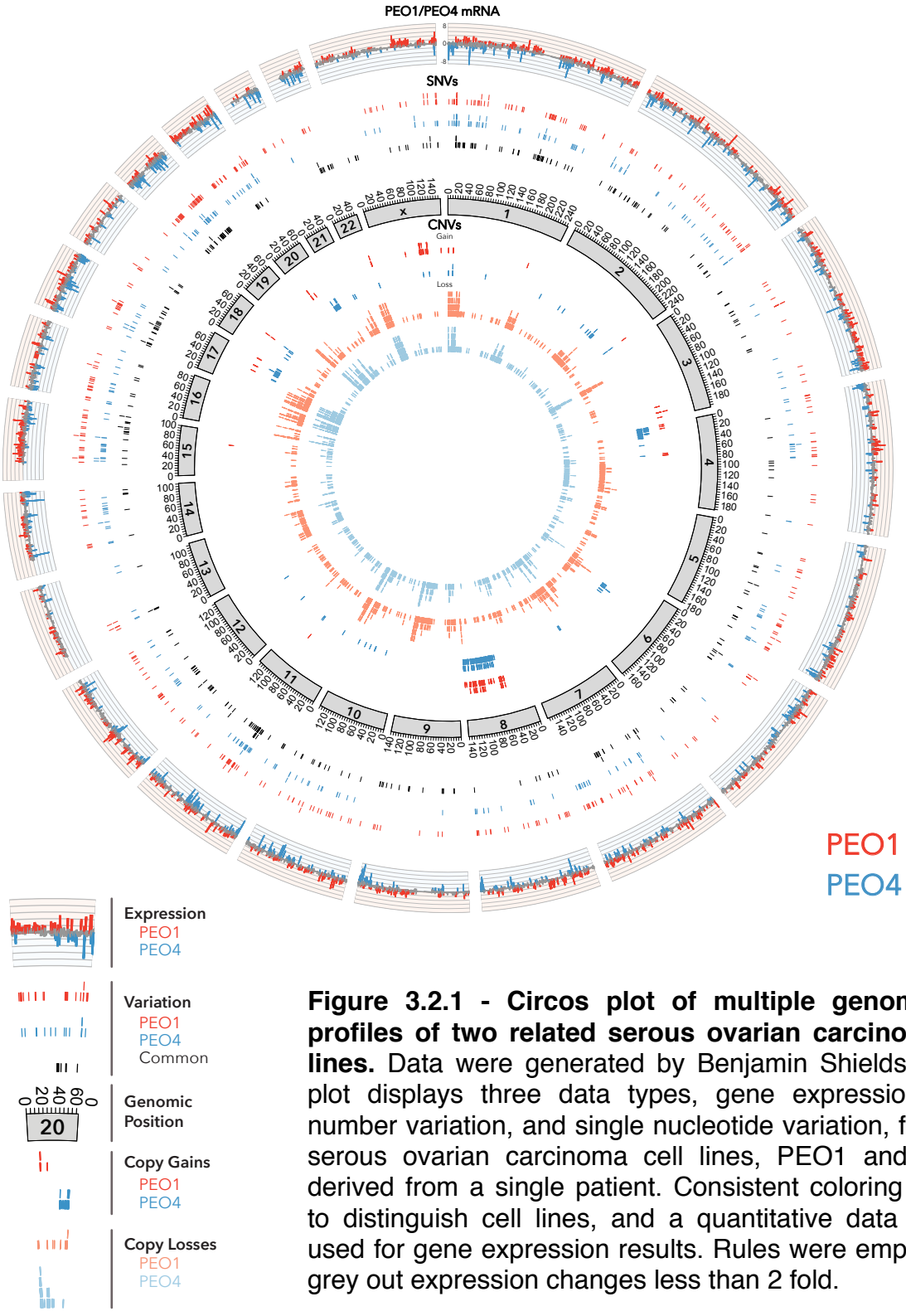


Figure 3.2.1 - Circos plot of multiple genome-wide profiles of two related serous ovarian carcinoma cell lines. Data were generated by Benjamin Shields. Circos plot displays three data types, gene expression, copy number variation, and single nucleotide variation, from two serous ovarian carcinoma cell lines, PEO1 and PEO4, derived from a single patient. Consistent coloring is used to distinguish cell lines, and a quantitative data track is used for gene expression results. Rules were employed to grey out expression changes less than 2 fold.

3.2.3 Conclusions

After several iterations over the display of three genome-wide data profiles from two cell lines, a single figure was generated in a manner that conveyed the significance and importance of the generated data. Over the course of revision and refinement an appreciation for the challenge of displaying the dynamic range of a dataset developed. Some data forms, such as variation, can be sufficiently represented via a presence/ absence demarcation with consistent coloring a marker of cell line context. Ratios such as relative gene expression between a pair of cell lines allow for the use of a two-dimensional axis and colored contexts to efficiently convey the direction and magnitude of the delta. The copy number variation data set contained both magnitude, direction and contextual data, and this presented several challenges. A color scale for the magnitude of copy number change that did not include white, and therefore risk having low or intermediate values merge with the background color, was not feasible without deviating from the consistent cell line coloring throughout the rest of the figure. Therefore, it was decided to independently mark gains and losses such that loss and gains across one cell line could be easily differentiated, and that distinctions between cell lines would be clear. While this did necessitate a dismissal of copy number data beyond the gain or loss relative to a normal tissue control, the graphic does allow for clear discrimination of copy number gains and losses both in a cell line and across the pair of cell lines.

3.3 SeedSync - A Method For Identifying miRNA Seed Co-occurrence

3.3.1 Introduction

The process of miRNA regulation of mRNA targets is regulated on several levels, including availability of individual miRNAs and mRNAs, sponge-like miRNA regulatory 3' UTRs, and shortening of 3' UTRs in proliferating cells [147,169]. miRNAs with shared seed sequences from different genomic loci can have redundant roles in regulating the expression of specific target genes, as is the case with the miR-25/92a family of miRNAs [122].

After observing the significant correlation between the toxicities of the miR-92a and miR-1226*, I was interested in determining if their was an enrichment for the co-occurrence of their seed sequences in the 3'UTRs. This would suggest that just as miRNAs and miRNA target sites have evolved both in and across species in a highly conserved manner, 3' UTRs may contain conserved pairings of miRNA target sites that allow for programmatic co-regulation of gene expression through multiple miRNAs independent of a sequence relationship.

3.3.2 Methodology

The statement of the essential question I intended to address here is fairly straightforward: Do any pairs miRNA target motifs co-occur more frequently than one would expect? My approach to this involves scanning annotated human 3' UTRs and tallying every case where a given pair of motifs co-occur. Instead of relying on the predictions of one of the myriad target prediction methods, I used the 7-mer motif

corresponding to the canonical seed sequence as the criteria for a miRNA target site. I then performed a search for each motif pair across all 3' UTRs, and as previously mentioned, tallied up the frequency of co-occurrence (Figure 3.3.1). This was performed via a parallelized Perl script, "SeedSync", with a FASTA 3' UTR file and FASTA file of mature miRNAs as the input files (Appendix C1).

However, a solitary observation of the occurrence of miRNA target sites in a single is insufficient for a salient observation; a measure of the meaningfulness of the observation is required. In this case I chose to determine the significance of the observed results empirically: the target space I was investigating, 3' UTRs, were randomized in a manner that preserved the nucleotide frequency of the original 3' UTR (Appendix C2). 1000 shuffled 3' UTR variants were then processed via SeedSync, and the output of these permutations was then collated into a single, flat file for downstream analysis and determination of significantly enriched motifs (Appendix C3). This allowed for the determination of an empirical p-value based on the frequency of tallies greater than the co-occurrence in the biological 3' UTR.

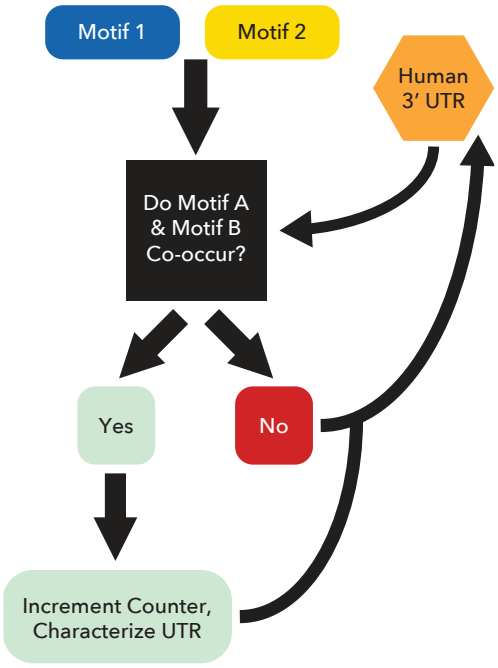


Figure 3.3.1 - Schematic of SeedSync workflow. Diagram graphically explaining the workflow for computational determination of the co-occurrence of two motifs in human 3' UTRs.

3.3.2 Results To Date

To date I have completed the analysis of the the hg19 human 3‘UTRs and 1,000 permuted UTRs for 1,203,176 unique pairs of 7-mer motifs. The median co-occurrence of a 7-mer pair in the 39,434 3‘UTRs was 654 (1.6%) with a standard deviation of 496 and minimum and maximum values of 0 and 6429, respectively (Figure 3.3.2 A). After calculating and tallying co-occurrence in permuted 3‘UTR collections I observed a surprising trend in the frequency of co-occurrence in the 1,000 permuted 3’ UTR collections. Specifically, 752,222 motif pairs had a co-occurrence frequency in the permuted 3’ UTRs of less than 50 (corresponding to an empirical p-value of 0.05), of which 736,013 had no observed co-occurrence values greater than the observed value ($p <= 0.001$) (Figure 3.3.2 B). An additional 386,340 pairs exceeded the observed value in the permuted 3’ UTRs greater than 950 times (Figure 3.3.2 B). The motifs corresponding to miR-92a, ATTGCAC, and miR-1226*, TGAGGGC, were observed to co-occur in 413 3’ UTRs, with an empirical p-value of less than 0.001. The observed co-occurrence is below the median of all motif pairs, and while the p-value suggests significance, I will describe why it is still difficult to determine whether or not this is a truly meaningful result.

These initial observations immediately suggests both that 7-mer co-occurrence may not be a random event, with certain miRNA pairs acting independently, and therefore having their motif co-occurrence be a very random event (high frequency of co-occurrence in the permuted data set) and others where the co-occurrence is not observed in the

permuted data set, suggesting that their co-occurrence suggests conserved co-regulation of shared target genes. However, refinements to the SeedSync method may be required to bolster these observations. The high number of motif pairs with no motif co-occurrence greater than the observed value in the permuted set makes it impossible to identify a set of meaningful co-occurrence for follow-up by empirical p-value. This could be addressed by increasing the permutation count; however, the calculations for the set of 1,000 permuted 3‘UTR collections took approximately 12,000 hours of clock time to compute. Significant improvements to the speed and efficiency of the algorithm in order to add a meaningful number of observations to this collection and would likely require a rewrite in a non-interpreted programming language such as C. Additional improvements include assessing the spacing between the observed motifs in order to rule out biologically implausible co-occurrence locations, as well as modeling the expected frequency of 7-mer occurrence in a sequence based on length and nucleotide frequency in a given sequence.

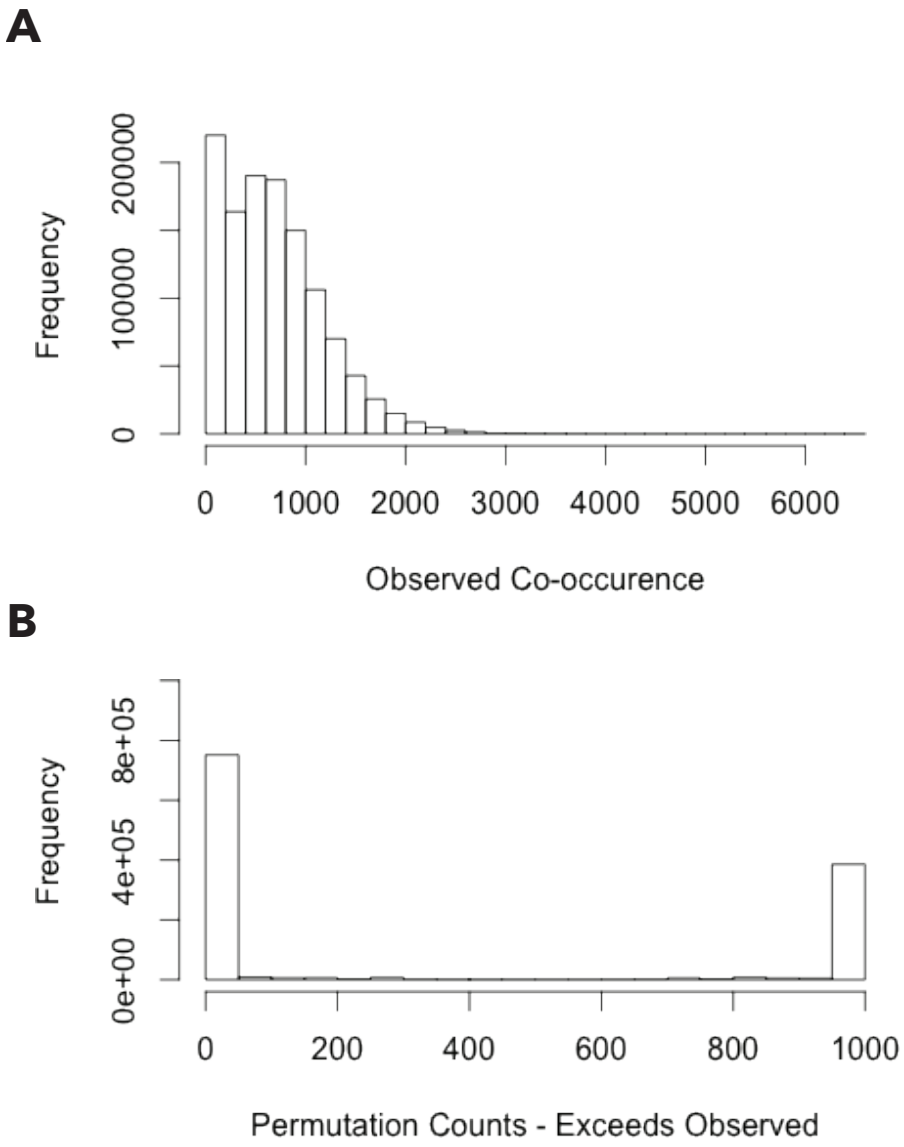


Figure 3.3.2 - Summary results of SeedSync analysis of the actual and permuted human 3'UTRs. A) Frequency distribution of co-occurrence of 7-mer motifs in human 3'UTRs. **B)** Frequency distribution of co-occurrences observed in the 1,000 permuted 3'UTRs that exceeded the value observed in the true 3' UTR set.

3.3.4 Future Work

After completing the additional work described above that will allow for the identification of significantly enriched motif pairs, work will commence on identifying biological contexts where the co-regulation of 3' UTRs by multiple miRNAs may be evident. The first step in this process will be observational and involve existing, freely available data sets containing both mRNA and miRNA profiles for the available samples. The TCGA has many data sets that fit these criteria. This is necessary as a first approach as it is unlikely that specific experiments have been published with miRNA and mRNA quantification after perturbation with two specific miRNAs of interest. By using large profile collections such as TCGA we can observe the relationship between mRNA abundance and high levels of a miRNA containing motif A, high levels of a miRNA motif B, or high levels of both miRNAs. I would expect to see the lowest gene expression in samples where both miRNAs were present at high levels, which would at least suggest that both miRNAs can regulate genes with 3' UTRs containing motifs from both miRNAs. More compelling evidence for co-regulation at the biological process level would involve combining observational data as I just described with gene set analyses that would demonstrate that genes with co-occurring motifs in the 3' UTR also share a common biological process.

Chapter 4: Conclusions and Future Directions

4.1 Discussion

Over this course of study I have identified a novel miRNA activity where the non-canonical activity of LNA miRNA inhibitors to miR-92a and miR-1226* selectively induce cell death by mimicking a 1α, 25-dihydroxyvitamin D_3 in a p53-depleted context. These inhibitors were identified via a miRNA inhibitor screen designed to identify oligos with selective toxicity between two histologically similar but genetically distinct cell lines, and the screen converged on this specific phenotype, as I established by showing the significant correlation between the miR-92a and miR-1226* inhibitors' toxicities. After observing a correlation between toxicity and low p53 mRNA and protein, I was able to show that loss of p53 in a resistant HBEC cell line was sufficient to sensitize the cells to the miR-92a inhibitor. Then, I observed that an RNA-based inhibitor did not reproduce the toxicity the miR-92a & miR-1226* inhibitors exhibited, and that the toxicity of the inhibitors was incredibly sequence specific. I then noticed that the single stranded miRNA inhibitors down-regulated the primary transcript and some of the mature miRNAs on the transcript, and that this effect was specific to the toxic inhibitors. Additionally, there was no threshold of primary transcript expression that conferred sensitivity to depletion. This suggested, and I confirmed by transcriptional profiling, that toxicity associated with the repression of the miR-17~92 transcript is context specific. I was finally able to demonstrate that the sensitivity to the miR-92a inhibitor predicted sensitivity of tumor cell lines to 1α, 25-dihydroxyvitamin D_3, a compound where defects

in its cellular metabolism are associated with poor prognosis in NSCLC, and that the miR-92a inhibitor created a 1α, 25-dihydroxyvitamin D_3-like response in sensitive cells.

The observation of a miRNA acting through a network of genes and in a context-dependent manner reaffirms expectations from some of my early days studying miRNA biology. The quantity of predicted miRNA targets for any given miRNA strongly suggests both that many different regulatory interactions “can” occur, as well as suggesting that circumstances where a one miRNA/one gene/one phenotype model holds true in a biological context are minimal. Additionally, a miRNA cannot regulate a target that is not present, and a miRNA would need to exert a solitary or dominant effect on an mRNA for its inhibition to have any meaningful impact on its abundance. These observations argue for additional systems-based approaches to understanding, testing, and conveying miRNA regulatory activity *in vitro* and *in vivo*.

Overall, even after the many years, investigations and dollars invested into the study of miRNAs there remains the potential for new and insightful observations. While the first miRNA mimic and inhibitor treatment based therapies are in development or clinical trials, the use of miRNA-based therapies as a component of combinatorial chemotherapies or in cases where adverse side effects prevent clinical use of a therapy remains largely unexplored. The results of this study demonstrate that the many-to-many relationship between miRNAs and their targets will need to be accounted for when identifying combinatorial or therapeutic-like activity of miRNAs in cancers. This study also demonstrated the ability of single-miRNA inhibition to identify miRNA losses

synthetic lethal with specific genetic changes in a particular tumor type. With the majority of NSCLC cases still lacking effective, targeted therapies additional studies similar to the one performed here will continue to identify synthetic lethal miRNA inhibitions that lead to more or better targeted therapeutic avenues in NSCLC and other cancers.

4.2 Future Works

There are several avenues for investigation that I believe would add additional depth and context to the investigations presented here. As the long-term ambition for this investigation was identifying miRNAs with potential therapeutic potential, I am conducting orthotopic mouse xenograft experiments with the miR-92a inhibitor and H358 cells. I am allowing tumor xenografts seeded by H358 cells to form in mouse flanks, and then administering the miR-92a inhibitor subcutaneously in order to arrest or reverse tumor growth. Tumor size, miR-92a abundance and histological assays for cell death will be used as endpoint assays for this study. As one of the major roadblocks between 1α, 25-dihydroxyvitamin D_3 and clinical relevance is hypercalcemia-related side effects at clinically relevant doses, this may represent a way to take advantage of that vulnerability in NSCLC while avoiding the unwanted toxicity.

Related to the 1α, 25-dihydroxyvitamin D_3 investigation, I have noticed that after treatment of the sensitive HBEC30-sh*TP53* with the miR-92a inhibitor much of the gene movement reflected what literature suggests is a keratinization/keratinocyte differentiation response. Keratinization is a process that is known to be driven by 1α,

25-dihydroxyvitamin D_3 and I was surprised to see it occurring in the bronchial epithelium [139,170,171]. Preliminary observations show substantial increases in the expression of various keratins after miR-92a treatment, as well as a host of extracellular matrix-resident structural and secreted proteins that may be related to the keratinization and differentiation process. Keratinocytes act as a barrier to the outside world as both a physical barrier to pathogens and insults outside the body, as well as by secreting immune factors and initiating a cutaneous inflammation response when necessary, and the terminal step in the differentiation process is in fact programmed cell death [170]. Further investigation into the strength of this differentiation signal in the lung epithelial lineage, and how it may become a vulnerability in neoplastic and malignant cells in that lineage would yield new insights on a potential point of vulnerability in NSCLC. This effort will require substantial investigations of sensitivity to keratinization signals in lung epithelium and transformed NSCLC lines, as well as the role of p53 in inhibiting this response in lung epithelium.

Finally, while I was able to demonstrate that down-regulation of the miR-17~92 transcript by a miR-92a inhibitor led to a de-repression of the miR-17~92 targetome in a context-dependent manner, there are still some questions of interest that could be addressed in future work. One relates to the two highly correlated miRNA inhibitors, miR-92a and miR-1226*. It would be an interesting study in the effect of sequence composition and motif structure to identify both what the minimally required sequence is to achieve the observed toxicity, as well as what shared motif is driving the similarity of the phenotypes between the two inhibitors. This could be accomplished through

meticulous manipulation of the oligo sequences. Additionally, if the methods progress to a point where the results could be cleanly interpreted it would be of interest to attempt to pull down the toxic oligos and identify the sequences to which the oligos are associated with. Results from this work could inform which motifs in the oligo are most important and inform the oligo study, and vice versa.

Appendices

Appendix A - Identifying high-confidence miR-17~92 miRNA targets

TargetScan conserved target DB retrieved June, 2013. Written for R 3.0.1.

```
conserved.targets <- read.table("~/Desktop/HBEC30 GeneExpression_032013/
analysis/target counting//Conserved_Site_Context_Scores.txt", sep = "\t",
header=TRUE)

mir1792 <- c("hsa-miR-17", "hsa-miR-18a", "hsa-miR-19a", "hsa-miR-20a", "hsa-
miR-19b", "hsa-miR-92a")

miRcluster.alltargets <- subset(conserved.targets, miRNA %in% mir1792)
miRcluster.alltargets.uniquegenes <- unique(miRcluster.alltargets
$Gene.Symbol)

miRcluster.alltargets.contextfilter75ile <- subset(miRcluster.alltargets,
context..score.percentile >= 75)
miRcluster.alltargets.contextfilter75ile.uniquegenes <-
unique(miRcluster.alltargets.contextfilter75ile$Gene.Symbol)

targets.df <- cbind(miRcluster.alltargets.uniquegenes,
miRcluster.alltargets.contextfilter75ile.uniquegenes)

gx <- read.table("gx.txt", sep="\t", header=TRUE)

gx.alltargets <- subset(gx, Name %in% miRcluster.alltargets.uniquegenes)
gx.contextfilter75ile <- subset(gx, Name %in%
miRcluster.alltargets.contextfilter75ile.uniquegenes)

write.table(gx.alltargets, file="gx.alltargets.txt", sep="\t")
write.table(gx.contextfilter75ile, file="gx.contextfiltered.txt", sep="\t")
```

Appendix B - Circos Plot Code

Code for the generation of Figure 3.x .Data generated by Benjamin Shields. Code is run by feeding this configuration file to the Circos v0.63 perl program.

```
<<include etc/colors_fonts_patterns.conf>>

<<include ideogram.conf>>
<<include ticks.conf>>

<image>
<<include etc/image.conf>>
</image>

karyotype = data/karyotype/karyotype.human.hg18.ben.txt

chromosomes_units = 1000000
chromosomes = -hsY
chromosomes_display_default = yes

<plots>

#PEO1 SNP Track
<plot>

type        = tile
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/
peo1.private.snv

r1 = 1.40r
r0 = 1.35r

layers_overflow=collapse
layers = 2
margin = 0.01u
orientation = out

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = reds-9-seq-6
</plot>

#PEO4 SNP Track
<plot>

type        = tile
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/
peo4.private.snv

r1 = 1.32r
r0 = 1.27r

layers_overflow=collapse
layers = 2
margin = 0.01u
```

```
orientation = out

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = blues-9-seq-6
</plot>

#Common SNP Track
<plot>

type        = tile
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/common.snv

r1 = 1.24r
r0 = 1.19r

layers_overflow=collapse
layers = 2
margin = 0.01u
orientation = out

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = black

</plot>

#PEO1 PEO4 Log2 RPKM+1 Ratio
<plot>

type        = histogram
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/log2ratio.txt

r1 = 1.65r
r0 = 1.5r
max = 8
min = -8

stroke_type = bin
thickness = 4
color = grey
fill_color  = grey
extend_bin = yes
skip_run    = yes

<backgrounds>
<background>
color       = grey
</background>

<background>
color       = grey
</background>
```

```
<background>
color       = reds-9-seq-1
y0                = 0.5r
</background>

<background>
color       = blues-9-seq-1
y0                = 0.0r
y1                = 0.5r
</background>

</backgrounds>

<axes>
<axis>
spacing = 0.1r
color = lgrey
thickness = 2
</axis>
</axes>


<rules>
<rule>
condition   = var(value) < -1
fill_color  = blues-9-seq-6
color       = blues-9-seq-6
</rule>

<rule>
condition   = var(value) > 1
fill_color  = reds-9-seq-6
color       = reds-9-seq-6
</rule>

<rule>
condition   = var(value) > -1
color = grey
fill_color  = grey
</rule>

<rule>
condition   = var(value) < 1
color = grey
fill_color  = grey
</rule>

</rules>

</plot>

##PEO1, PEO4 Gains CNV
<plot>
type        = tile
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/peo1.cnv.gain

r1 = 0.90r
r0 = 0.85r
```

```
layers = 3
margin = 0.02u
orientation = out
layers_overflow=collapse

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = reds-9-seq-6

</plot>

#PEO4 Gain CNV
<plot>
type        = tile
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/peo4.cnv.gain

r1 = 0.81r
r0 = 0.76r

layers = 3
margin = 0.02u
orientation = out
layers_overflow=collapse

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = blues-9-seq-6
</plot>

#PEO1, PEO4 Loss CNV
<plot>
type        = tile
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/peo1.cnv.loss

r1 = 0.70r
r0 = 0.60r

layers = 5
margin = 0.02u
orientation = out
layers_overflow=collapse

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = reds-9-seq-4

</plot>

#PEO4 Loss CNV
<plot>
type        = tile
```

```
file        = /Users/rborkows/Desktop/circos.work ben v3.1/data/peo4.cnv.loss

r1 = 0.56r
r0 = 0.46r

layers = 5
margin = 0.02u
orientation = out
layers_overflow=collapse

thickness = 15
padding = 5

stroke_thickness  = 2
stroke_color            = blues-9-seq-4
</plot>

</plots>

<<include etc/housekeeping.conf>>
```

Appendix C - SeedSync Data Processing and Analysis Code

1) Perl code for performing the SeedSync analysis.

```
#!/usr/bin/perl -w

#SeedSync, a utility to identify motifs (miRNA seeds) occuring in the same
UTR.
#syntax: seedsync.pl <FASTA sequence file> <FASTA mature miRNAs>

#MAKE SURE MOTIFS ARE IN ALL CAPS!

{ use Bio::SeqIO;
 use Math::Combinatorics;
 use Parallel::ForkManager;
      use List::MoreUtils qw(uniq);
      #use Statistics::Descriptive;
      #use Text::Match::FastAlternatives;
      use Time::HiRes qw(gettimeofday tv_interval); #get better than 1 second
resolution

 $MAX_PROCESSES=8; #probably don't want this to be greater than the # of
cores you have.

 #generate array of all possible miRNA seed combinations
 print "opening miRNA FASTA file...";
 $mirs = Bio::SeqIO->new(-file => "$ARGV[1]", -format => 'fasta');
 print "ok.\n";

 $record;
 @hsa_mir_SEQ;
      @hsa_mir_SEQ_temp;

 #slice out human seed sequences (bases 2-8)
 print "Pull out bases 2-n of miRNA sequence...";
 while( $record = $mirs->next_seq() ) {

 $seq=$record->seq;
 $id=$record->id;

 if($id =~ /hsa/) {
 $seed = substr $seq, 1, 7;
 #print "$seed\n";
 push(@hsa_mir_SEQ_temp, $seed);
 }
 }
 print "ok.\n Filtering duplicate seeds...\n";

      @hsa_mir_SEQ = uniq @hsa_mir_SEQ_temp;
      undef @hsa_mir_SEQ_temp;

 #array of all possible combinations
 print "generating all possible combinations...\n";
 @combi = combine(2, @hsa_mir_SEQ);
```

```
print "\n$#combi generated, let's clean out identical pairs...\n";
undef @hsa_mir_SEQ;

#this will leave motif pairs that are identical n-mers. Weed these out!

for $p (0 .. $#combi){
if("$combi[$p][0]" ne "$combi[$p][1]"){
@goodPair=($combi[$p][0],$combi[$p][1]);
#print "$goodPair[0], $goodPair[1]\n";
push(@combinations, [@goodPair]);
}

}

undef @combi;

print "\n$#combinations pairs remain.";

#parse 3' UTR FASTA file

$u;
@recs;
$utr_file = Bio::SeqIO->new(-file => "$ARGV[0]", -format => 'fasta' -
alphabet => 'dna');
while( $u = $utr_file->next_seq() ){
$string=$u->seq;
push(@recs,$string);
}
undef $utr_file;
$pm = new Parallel::ForkManager($MAX_PROCESSES);

open(OUTFILE, ">seedpairs_numhits.test.txt");
print OUTFILE "Motif 1\tMotif 2\tCount\n";

#for $i (0 .. $#combinations){
$tloop0 = [gettimeofday];
      for $i (0 .. 1000){
# Forks and returns the pid for the child:

$pm->start and next; # do the fork

           #Time each loop.
           $t0 = [gettimeofday];

$motif1 = $combinations[$i][0];
$motif2 = $combinations[$i][1];

$count=0;

print "processing motifs: $motif1\t$motif2\t";
for $j ( 0 .. $#recs ) {

$string = $recs[$j];
#search for both motifs in one UTR
if($string =~ /$motif1/ && $string =~ /$motif2/){
                  $count++;
}
}
```

```
 #list stats
            $t1 = [gettimeofday];
 print "$count\t";
            print tv_interval($t0,$t1);
            print "\n";
 print OUTFILE "$motif1\t$motif2\t$count\n";

 $pm->finish; # Terminates the child process
 }
 $pm->wait_all_children;

      $tloop1 = [gettimeofday];
      $looptime = tv_interval($tloop0,$tloop1);
 print "Total Loop time: $looptime\n";


 close OUTFILE;
 print "\nDone!\n";
}
```

2) Perl code for generating shuffled 3' UTRs.

```
#!/usr/bin/perl -w

#Shuffles all sequences in a FASTA file using Fisher-yates shuffle.
#Usage: perl seqshuffle.pl <FASTA file> <shuffle iterations>
#ex: 'seqshuffle.pl 3p_hsa_hg19 20' would create 20 shuffled variants of
"3p_hsa_hg19"

{     use strict;

      use Bio::SeqIO;
      use Array::Shuffle qw(shuffle_array);
      use Parallel::ForkManager;

      our $MAX_PROCESSES=8;

      our $file = $ARGV[0];
      our $iterations = $ARGV[1];
      mkdir ("output") || die "Can't create output directory\n";

      our $pm = new Parallel::ForkManager($MAX_PROCESSES);

      for my $p (1 .. $iterations){

            my $filename = join(".", $file, $p);

            open(OUTFILE, ">output/$filename") || die "Can't create output
file.\n";

            my $sequences = Bio::SeqIO->new(-file => $file, '-format' =>
'Fasta');

       print "$p\n";

            while( my $record = $sequences->next_seq() ) {

                  $pm->start and next; # do the fork

       my $seq=$record->seq;
                  my $id=$record->id;

                  my @sequence = split("", $seq);
                  shuffle_array(@sequence);
                  my $shuffled = join("", @sequence);

                  print OUTFILE ">$id\n$shuffled\n";

                  $pm->finish;
            }
            $pm->wait_all_children;

            close OUTFILE;
      }

}
```

3) Perl code for merging permutations into one flat-file for analysis.

```
#!/usr/bin/perl -w

{
 @motif1 = (); #ordered list from "Motif 1" column in permutation 1 results.
Motif 2 follows $
 @motif2 = ();
 %lookupHash = ();

 #Read in file1 to populate the list of records
 $file1 = "";
 open (DATA, "<seedpairs_output.hg19_utr.1") || die "Can't open first file.";
 while (<DATA>) {
 $file1 .= $_;
 }
 close DATA;

 @records = split/\n/, $file1; #split file1 into one array element per line
 shift(@records); #drop hearder

 $counter = 0;
 print "\n File read okay. Populating lookup hash.";

 foreach $f (@records) {
 chomp ($f);
 @f_fields = split /\t/, $f;
 push(@motif1, $f_fields[0]);
 push(@motif2, $f_fields[1]);
 $lookupHash{$f_fields[0] . $f_fields[1]} = $counter-1; #add records to
lookup hash w/ $
 $lookupHash{$f_fields[1] . $f_fields[0]} = $counter-1; #Counter-1 because of
zero-inde$
 print "$f_fields[0] $f_fields[1]\t$lookupHash{$f_fields[0].$f_fields[1]}\n";
 }
      print "\n hashes populated. Aggregating file contents...";
 @counts = ();

      for ($i = 1; $i <= 1000; $i ++){

            $file = "";
            $filename = "seedpairs_output.hg19_utr.".$i;
       open (DATA, "<".$filename) || die "Can't open $filename";
       while (<DATA>) {
       $file .= $_;
       }
       close DATA;

            foreach $c (@file){
       chomp ($c);
       @c_fields = split /\t/, $c;
                  $motifPos = $lookupHash{$c_fields[0].$c_fields[1]};
                  $counts[$motifPos][$i] = $c_fields[2]; #Please don't blow
up.
                  print "Stored count: $counts[$motifPos][$i]";
            }

      }
```

```
     print "\n aggregation complete, print output.";

 open (OUTFILE, ">merged_permutations.txt");

     print OUTFILE "Motif1\tMotif2\n";

     @filenames1000 = 1..1000;
 foreach $x (@filenames1000) {
 print OUTFILE "$x\t";
 }
 print OUTFILE "\n";


 for ($i = 0; $i <= $motif1; $i ++){
 print OUTFILE "$motif1[$i]\t$motif2[$i]";

 for($j = 0; $j <= $999; $j ++){

 print OUTFILE "\t$counts[$i][$j]";

 }



 print OUTFILE "\n";

 }



 close OUTFILE || die;

}
```

References

1. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2009 Incidence and Mortality Web-based Report. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2013. Available at: www.cdc.gov/uscs. Accessed 23 May 2013.

2. Siegel R, Naishadham D, Jemal A (2013) Cancer statistics, 2013. CA Cancer J Clin 63: 11–30. doi:10.3322/caac.21166.

3. Hecht SS (1999) Tobacco smoke carcinogens and lung cancer. J Natl Cancer Inst 91: 1194–1210.

4. Office of the Surgeon General (1964) Smoking and health. Report of the Advisory Committee to the Surgeon General of the Public Health Service. US Gov Print Off: 1–386.

5. Group N-SCLCC (1995) Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. Non-small Cell Lung Cancer Collaborative Group. BMJ 311: 899–909.

6. Heist RS, Engelman JA (2012) SnapShot: Non-Small Cell Lung Cancer. Cancer Cell 21: 448–448.e2. doi:10.1016/j.ccr.2012.03.007.

7. Winton T, Livingston R, Johnson D, Rigas J, Johnston M, et al. (2005) Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. N Engl J Med 352: 2589–2597. doi:10.1056/NEJMoa043623.

8. Pujol J-L, Barlesi F, Daurès J-P (2006) Should chemotherapy combinations for advanced non-small cell lung cancer be platinum-based? A meta-analysis of phase III randomized trials. Lung Cancer 51: 335–345. doi:10.1016/j.lungcan. 2005.11.001.

9. Douillard J-Y, Rosell R, De Lena M, Carpagnano F, Ramlau R, et al. (2006) Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB-IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. The lancet oncology 7: 719–727. doi:10.1016/S1470-2045(06)70804-X.

10. Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon J-P, et al. (2004) Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. N Engl J Med 350: 351–360. doi:10.1056/ NEJMoa031644.

11. D'Addario G, Pintilie M, Leighl NB, Feld R, Cerny T, et al. (2005) Platinum-based versus non-platinum-based chemotherapy in advanced non-small-cell lung cancer: a meta-analysis of the published literature. J Clin Oncol 23: 2926–2936. doi:10.1200/JCO.2005.03.045.

12. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100: 57–70.

13. Stehelin D, Varmus HE, Bishop JM, Vogt PK (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 260: 170–173.

14. Roock WD, Vriendt VD, Normanno N, Ciardiello F, Tejpar S (2010) KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. The lancet oncology. doi:10.1016/S1470-2045(10)70209-6.

15. Riely GJ, Marks J, Pao W (2009) KRAS mutations in non-small cell lung cancer. Proceedings of the American Thoracic Society 6: 201–205. doi:10.1513/pats. 200809-107LC.

16. Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, et al. (2005) A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res 65: 9628–9632. doi: 10.1158/0008-5472.CAN-05-2352.

17. Knudson AG (1971) Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci USA 68: 820–823.

18. Sherr CJ (2004) Principles of tumor suppression. Cell 116: 235–246. doi:10.1016/ S0092-8674(03)01075-4.

19. Baker SJ, Markowitz S, Fearon ER, Willson JK, Vogelstein B (1990) Suppression of human colorectal carcinoma cell growth by wild-type p53. Science 249: 912–915.

20. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531–537. doi:10.1126/science.286.5439.531.

21. Schmutz J, Wheeler J, Grimwood J, Dickson M, Yang J, et al. (2004) Quality assessment of the human genome sequence. Nature 429: 365–368. doi:10.1038/nature02390.

22. Singh A, Greninger P, Rhodes D, Koopman L, Violette S, et al. (2009) A Gene Expression Signature Associated with "K-Ras Addiction" Reveals Regulators of EMT and Tumor Cell Survival. Cancer Cell 15: 489–500. doi:10.1016/j.ccr.2009.03.022.

23. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, et al. (2000) Molecular portraits of human breast tumours : Article : Nature. Nature 406: 747–752. doi:10.1038/35021093.

24. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, et al. (2000) Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 24: 227–235. doi:10.1038/73432.

25. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, et al. (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503–511. doi:10.1038/35000501.

26. Whitehurst AW, Bodemann BO, Cardenas J, Ferguson D, Girard L, et al. (2007) Synthetic lethal screen identification of chemosensitizer loci in cancer cells. Nature 446: 815–819. doi:10.1038/nature05697.

27. Bamford S, Dawson E, Forbes S, Clements J, Pettett R, et al. (2004) The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. Br J Cancer 91: 355–358. doi:10.1038/sj.bjc.6601894.

28. Forbes S, Clements J, Dawson E, Bamford S, Webb T, et al. (2006) COSMIC 2005. Br J Cancer 94: 318–322. doi:10.1038/sj.bjc.6602928.

29. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, et al. (2008) The Catalogue of Somatic Mutations in Cancer (COSMIC). Current protocols in human genetics / editorial board, Jonathan L Haines [et al] Chapter 10: Unit10.11. doi: 10.1002/0471142905.hg1011s57.

30. Ding L, Getz G, Wheeler DA, Mardis ER, Mclellan MD, et al. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455: 1069–1075. doi:10.1038/nature07423.

31. Pao W, Girard N (2011) New driver mutations in non-small-cell lung cancer. The lancet oncology 12: 175–180. doi:10.1016/S1470-2045(10)70087-5.

32. Weinstein IB (2002) Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science 297: 63–64. doi:10.1126/science.1073096.

33. Fisher GH, Wellen SL, Klimstra D, Lenczowski JM, Tichelaar JW, et al. (2001) Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. Genes Dev 15: 3249–3262. doi:10.1101/gad.947701.

34. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144: 646–674. doi:10.1016/j.cell.2011.02.013.

35. Eberhard DA, Johnson BE, Amler LC, Goddard AD, Heldens SL, et al. (2005) Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. J Clin Oncol 23: 5900–5909. doi:10.1200/JCO.2005.02.857.

36. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, et al. (2004) EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 304: 1497–1500. doi:10.1126/science.1099314.

37. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, et al. (2011) Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med 3: 75ra26. doi:10.1126/scitranslmed.3002003.

38. Turke AB, Zejnullahu K, Wu Y-L, Song Y, Dias-Santagata D, et al. (2010) Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell 17: 77–88. doi:10.1016/j.ccr.2009.11.022.

39. Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, et al. (2010) Inferring tumor progression from genomic heterogeneity. Genome Res 20: 68–80. doi:10.1101/gr.099622.109.

40. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90–94. doi:10.1038/nature09807.

41. Bendall SC, Simonds EF, Qiu P, Amir EAD, Krutzik PO, et al. (2011) Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. Science 332: 687–696. doi:10.1126/science.1198704.

42. Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, et al. (2012) Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. Nat Biotech. doi:10.1038/nbt.2317.

43. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75: 843–854.

44. Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP (2003) Vertebrate microRNA genes. Science 299: 1540. doi:10.1126/science.1080372.

45. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294: 853–858. doi:10.1126/science.1064921.

46. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294: 858–862. doi:10.1126/science.1065062.

47. Lee RC, Ambros V (2001) An extensive class of small RNAs in Caenorhabditis elegans. Science 294: 862–864. doi:10.1126/science.1065329.

48. Moss EG, Poethig RS (2002) MicroRNAs: something new under the sun. Curr Biol 12: R688–R690.

49. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297.

50. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, et al. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403: 901–906. doi:10.1038/35002607.

51. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, et al. (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408: 86–89. doi:10.1038/35040556.

52. Hutvágner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, et al. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293: 834–838. doi:10.1126/science.1062961.

53. Lee Y, Kim M, Han J, Yeom K-H, Lee S, et al. (2004) MicroRNA genes are transcribed by RNA polymerase II. EMBO J 23: 4051–4060. doi:10.1038/sj.emboj.7600385.

54. Berezikov E, Chung W-J, Willis J, Cuppen E, Lai EC (2007) Mammalian mirtron genes. Mol Cell 28: 328–336. doi:10.1016/j.molcel.2007.09.028.

55. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT (2005) c-Myc-regulated microRNAs modulate E2F1 expression. Nature 435: 839–843. doi:10.1038/nature03677.

56. Lee Y, Jeon K, Lee J-T, Kim S, Kim VN (2002) MicroRNA maturation: stepwise processing and subcellular localization. EMBO J 21: 4663–4670.

57. Lee Y, Ahn C, Han J, Choi H, Kim J, et al. (2003) The nuclear RNase III Drosha initiates microRNA processing. Nature 425: 415–419. doi:10.1038/nature01957.

58. Yi R, Qin Y, Macara IG, Cullen BR (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 17: 3011–3016. doi:10.1101/gad.1158803.

59. Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409: 363–366. doi:10.1038/35053110.

60. Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell 123: 607–620. doi:10.1016/j.cell.2005.08.044.

61. Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. Cell 115: 209–216.

62. Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, et al. (2003) Asymmetry in the assembly of the RNAi enzyme complex. Cell 115: 199–208.

63. Rand T, Petersen S, Du F, Wang X (2005) Argonaute2 Cleaves the Anti-Guide Strand of siRNA during RISC Activation. Cell 123: 621–629. doi:10.1016/j.cell.2005.10.020.

64. Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, et al. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. Science 305: 1437–1441. doi:10.1126/science.1102513.

65. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R (2005) Human RISC Couples MicroRNA Biogenesis and Posttranscriptional Gene Silencing. Cell 123: 631–640. doi:10.1016/j.cell.2005.10.022.

66. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA (2008) In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci USA 105: 512–517. doi:10.1073/pnas.0710869105.

67. Rana TM (2007) Illuminating the silence: understanding the structure and function of small RNAs. Nat Rev Mol Cell Biol 8: 23–36. doi:10.1038/nrm2085.

68. Guo H, Ingolia NT, Weissman JS, Bartel DP (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466: 835–840. doi: 10.1038/nature09267.

69. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215–233. doi:10.1016/j.cell.2009.01.002.

70. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15–20. doi:10.1016/j.cell.2004.12.035.

71. Lewis BP, Shih I-H, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115: 787–798.

72. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, et al. (2005) Combinatorial microRNA target predictions. Nat Genet 37: 495–500. doi:10.1038/ng1536.

73. Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3: e85. doi:10.1371/journal.pbio.0030085.

74. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, et al. (2002) Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 99: 15524–15529. doi:10.1073/ pnas.242606799.

75. Kumar MS, Pester RE, Chen CY, Lane K, Chin C, et al. (2009) Dicer1 functions as a haploinsufficient tumor suppressor. Genes Dev 23: 2700–2704. doi:10.1101/gad.1848209.

76. Volinia S, Calin GA, Liu C-G, Ambs S, Cimmino A, et al. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci USA 103: 2257–2261. doi:10.1073/pnas.0510565103.

77. Hatley ME, Patrick DM, Garcia MR, Richardson JA, Bassel-Duby R, et al. (2010) Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21. Cancer Cell 18: 282–293. doi:10.1016/j.ccr.2010.08.013.

78. Chan JA, Krichevsky AM, Kosik KS (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res 65: 6029–6033. doi:10.1158/0008-5472.CAN-05-0137.

79. Ciafrè SA, Galardi S, Mangiola A, Ferracin M, Liu C-G, et al. (2005) Extensive modulation of a set of microRNAs in primary glioblastoma. Biochem Biophys Res Commun 334: 1351–1358. doi:10.1016/j.bbrc.2005.07.030.

80. Tran N, McLean T, Zhang X, Zhao CJ, Thomson JM, et al. (2007) MicroRNA expression profiles in head and neck cancer cell lines. Biochem Biophys Res Commun 358: 12–17. doi:10.1016/j.bbrc.2007.03.201.

81. Iorio MV, Visone R, Di Leva G, Donati V, Petrocca F, et al. (2007) MicroRNA signatures in human ovarian cancer. Cancer Res 67: 8699–8707. doi:10.1158/0008-5472.CAN-07-1936.

82. Lawrie CH, Soneji S, Marafioti T, Cooper CDO, Palazzo S, et al. (2007) MicroRNA expression distinguishes between germinal center B cell-like and activated B cell-like subtypes of diffuse large B cell lymphoma. Int J Cancer 121: 1156–1161. doi: 10.1002/ijc.22800.

83. Lui W-O, Pourmand N, Patterson BK, Fire A (2007) Patterns of known and novel small RNAs in human cervical cancer. Cancer Res 67: 6031–6043. doi: 10.1158/0008-5472.CAN-06-0561.

84. Meng F, Henson R, Wehbe-Janek H, Ghoshal K, Jacob ST, et al. (2007) MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. Gastroenterology 133: 647–658. doi:10.1053/ j.gastro.2007.05.022.

85. Weiss GJ, Bemis LT, Nakajima E, Sugita M, Birks DK, et al. (2008) EGFR regulation by microRNA in lung cancer: correlation with clinical response and survival to gefitinib and EGFR expression in cell lines. Ann Oncol 19: 1053–1059. doi: 10.1093/annonc/mdn006.

86. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, et al. (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci USA 102: 13944–13949. doi:10.1073/pnas.0506654102.

87. Bandi N, Zbinden S, Gugger M, Arnold M, Kocher V, et al. (2009) miR-15a and miR-16 are implicated in cell cycle regulation in a Rb-dependent manner and are

frequently deleted or down-regulated in non-small cell lung cancer. Cancer Res 69: 5553–5559. doi:10.1158/0008-5472.CAN-08-4277.

88. He L, He X, Lowe SW, Hannon GJ (2007) microRNAs join the p53 network--another piece in the tumour-suppression puzzle. Nat Rev Cancer 7: 819–822. doi: 10.1038/nrc2232.

89. Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, et al. (2007) p53-mediated activation of miRNA34 candidate tumor-suppressor genes. Curr Biol 17: 1298–1307. doi:10.1016/j.cub.2007.06.068.

90. Du L, Pertsemlidis A (2010) microRNAs and lung cancer: tumors and 22-mers. Cancer Metastasis Rev 29: 109–122. doi:10.1007/s10555-010-9204-9.

91. He L, He X, Lim LP, de Stanchina E, Xuan Z, et al. (2007) A microRNA component of the p53 tumour suppressor network. Nature 447: 1130–1134. doi:10.1038/ nature05939.

92. Hermeking H (2010) The miR-34 family in cancer and apoptosis. Cell Death Differ 17: 193–199. doi:10.1038/cdd.2009.56.

93. Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, et al. (2007) Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. Cell Cycle 6: 1586–1593.

94. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, et al. (2005) RAS is regulated by the let-7 microRNA family. Cell 120: 635–647. doi:10.1016/j.cell.2005.01.014.

95. Chin LJ, Ratner E, Leng S, Zhai R, Nallur S, et al. (2008) A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. Cancer Res 68: 8535–8540. doi:10.1158/0008-5472.CAN-08-2129.

96. Lee YS, Dutta A (2007) The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev 21: 1025–1030. doi:10.1101/gad.1540407.

97. Gibbons DL, Lin W, Creighton CJ, Rizvi ZH, Gregory PA, et al. (2009) Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. Genes Dev 23: 2140–2151. doi:10.1101/gad.1820209.

98. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, et al. (2008) A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. EMBO Rep 9: 582–589. doi:10.1038/embor.2008.74.

99. Park S-M, Gaur AB, Lengyel E, Peter ME (2008) The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev 22: 894–907. doi:10.1101/gad.1640608.

100. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, et al. (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol 10: 593–601. doi:10.1038/ncb1722.

101. Korpal M, Lee ES, Hu G, Kang Y (2008) The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. J Biol Chem 283: 14910–14914. doi:10.1074/jbc.C800074200.

102. Ding W, Dang H, You H, Steinway S, Takahashi Y, et al. (2012) miR-200b restoration and DNA methyltransferase inhibitor block lung metastasis of mesenchymal-phenotype hepatocellular carcinoma. Oncogenesis 1: e15. doi: 10.1038/oncsis.2012.15.

103. Du L, Subauste MC, DeSevo C, Zhao Z, Baker M, et al. (2012) miR-337-3p and its targets STAT3 and RAP1A modulate taxane sensitivity in non-small cell lung cancers. PLoS ONE 7: e39167. doi:10.1371/journal.pone.0039167.

104. Wiggins JF, Ruffino L, Kelnar K, Omotola M, Patrawala L, et al. (2010) Development of a lung cancer therapeutic based on the tumor suppressor microRNA-34. Cancer Res 70: 5923–5930. doi: 10.1158/0008-5472.CAN-10-0655.

105. Krützfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, et al. (2005) Silencing of microRNAs in vivo with 'antagomirs'. Nature 438: 685–689. doi:10.1038/nature04303.

106. Ma L, Reinhardt F, Pan E, Soutschek J, Bhat B, et al. (2010) Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. Nat Biotech 28: 341–347. doi:10.1038/nbt.1618.

107. Lanford RE, Hildebrandt-Eriksen ES, Petri A, Persson R, Lindow M, et al. (2010) Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. Science 327: 198–201. doi:10.1126/science.1178178.

108. Trang P, Medina PP, Wiggins JF, Ruffino L, Kelnar K, et al. (2010) Regression of murine lung tumors by the let-7 microRNA. Oncogene 29: 1580–1587. doi: 10.1038/onc.2009.445.

109. Trang P, Wiggins JF, Daige CL, Cho C, Omotola M, et al. (2011) Systemic Delivery of Tumor Suppressor microRNA Mimics Using a Neutral Lipid Emulsion Inhibits Lung Tumors in Mice. Molecular Therapy 19: 1116–1122. doi:10.1038/mt. 2011.48.

110. Stenvang J, Petri A, Lindow M, Obad S, Kauppinen S (2012) Inhibition of microRNA function by antimiR oligonucleotides. Silence 3: 1. doi: 10.1186/1758-907X-3-1.

111. Anand S, Majeti BK, Acevedo LM, Murphy EA, Mukthavaram R, et al. (2010) MicroRNA-132-mediated loss of p120RasGAP activates the endothelium to facilitate pathological angiogenesis. Nat Med 16: 909–914. doi:10.1038/nm.2186.

112. Obad S, Santos dos CO, Petri A, Heidenblad M, Broom O, et al. (2011) Silencing of microRNA families by seed-targeting tiny LNAs. Nat Genet 43: 371–378. doi: 10.1038/ng.786.

113. Zhou C, Wu Y-L, Chen G, Feng J, Liu X-Q, et al. (2011) Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. The lancet oncology 12: 735–742. doi: 10.1016/S1470-2045(11)70184-X.

114. Kwak EL, Bang Y-J, Camidge DR, Shaw AT, Solomon B, et al. (2010) Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 363: 1693–1703. doi:10.1056/NEJMoa1006448.

115. Bunn PA (2012) Worldwide overview of the current status of lung cancer diagnosis and treatment. Arch Pathol Lab Med 136: 1478–1481. doi:10.5858/arpa. 2012-0295-SA.

116. Liu L, Shao X, Gao W, Zhang Z, Liu P, et al. (2012) MicroRNA-133b inhibits the growth of non-small-cell lung cancer by targeting the epidermal growth factor receptor. FEBS J 279: 3800–3812. doi:10.1111/j.1742-4658.2012.08741.x.

117. MacFarlane L-A, R Murphy P (2010) MicroRNA: Biogenesis, Function and Role in Cancer. CG 11: 537–561. doi:10.2174/138920210793175895.

118. Kota J, Chivukula RR, O'Donnell KA, Wentzel EA, Montgomery CL, et al. (2009) Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model. Cell 137: 1005–1017. doi:10.1016/j.cell.2009.04.021.

119. Elmén J, Lindow M, Schütz S, Lawrence M, Petri A, et al. (2008) LNA-mediated microRNA silencing in non-human primates. Nature 452: 896–899. doi:10.1038/nature06783.

120. Torres AG, Fabani MM, Vigorito E, Gait MJ (2011) MicroRNA fate upon targeting with anti-miRNA oligonucleotides as revealed by an improved Northern-blot-based method for miRNA detection. RNA 17: 933–943. doi:10.1261/rna.2533811.

121. Liu J, Lee W, Jiang Z, Chen Z, Jhunjhunwala S, et al. (2012) Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. Genome Res 22: 2315–2327. doi:10.1101/gr.140988.112.

122. Petrocca F, Vecchione A, Croce CM (2008) Emerging Role of miR-106b-25/miR-17-92 Clusters in the Control of Transforming Growth Factor Signaling. Cancer Res 68: 8191–8194. doi:10.1158/0008-5472.CAN-08-1768.

123. Mendell JT (2008) miRiad roles for the miR-17-92 cluster in development and disease. Cell 133: 217–222. doi:10.1016/j.cell.2008.04.001.

124. Dowell JE, Minna JD (2006) EGFR mutations and molecularly targeted therapy: a new era in the treatment of lung cancer. Nat Clin Pract Oncol 3: 170–171. doi:10.1038/ncponc0476.

125. Weisz L, Damalas A, Liontos M, Karakaidos P, Fontemaggi G, et al. (2007) Mutant p53 enhances nuclear factor kappaB activation by tumor necrosis factor alpha in cancer cells. Cancer Res 67: 2396–2401. doi:10.1158/0008-5472.CAN-06-2425.

126. Stambolsky P, Tabach Y, Fontemaggi G, Weisz L, Maor-Aloni R, et al. (2010) Modulation of the vitamin D3 response by cancer-associated mutant p53. Cancer Cell 17: 273–285. doi:10.1016/j.ccr.2009.11.025.

127. Levine AJ (1997) p53, the cellular gatekeeper for growth and division. Cell 88: 323–331.

128. Schetter AJ, Heegaard NHH, Harris CC (2010) Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. Carcinogenesis 31: 37–49. doi:10.1093/carcin/bgp272.

129. Gudkov AV, Gurova KV, Komarova EA (2011) Inflammation and p53: A Tale of Two Stresses. Genes & Cancer 2: 503–516. doi:10.1177/1947601911409747.

130. Sato M, Larsen JE, Lee W, Sun H, Shames DS, et al. (2013) Human lung epithelial cells progressed to malignancy through specific oncogenic manipulations. Molecular Cancer Research. doi:10.1158/1541-7786.MCR-12-0634-T.

131. Lu Y, Thomson JM, Wong HYF, Hammond SM, Hogan BLM (2007) Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. Dev Biol 310: 442–453. doi:10.1016/j.ydbio.2007.08.007.

132. Yan H-L, Xue G, Mei Q, Wang Y-Z, Ding F-X, et al. (2009) Repression of the miR-17-92 cluster by p53 has an important function in hypoxia-induced apoptosis. EMBO J 28: 2719–2732. doi:10.1038/emboj.2009.214.

133. Nittner D, Lambertz I, Clermont F, Mestdagh P, Köhler C, et al. (2012) Synthetic lethality between Rb, p53 and Dicer or miR-17-92 in retinal progenitors suppresses retinoblastoma formation. Nat Cell Biol 14: 958–965. doi:10.1038/ncb2556.

134. Bail S, Swerdel M, Liu H, Jiao X, Goff LA, et al. (2010) Differential regulation of microRNA stability. RNA 16: 1032–1039. doi:10.1261/rna.1851510.

135. Winter J, Diederichs S (2011) Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due to microRNA stabilization. RNA Biol 8: 1149–1157. doi:10.4161/rna.8.6.17665.

136. DeLuca HF (2004) Overview of general physiologic features and functions of vitamin D. Am J Clin Nutr 80: 1689S–96S.

137. Müller K, Bendtzen K (1996) 1,25-Dihydroxyvitamin D3 as a natural regulator of human immune functions. J Investig Dermatol Symp Proc 1: 68–71.

138. Peehl DM, Skowronski RJ, Leung GK, Wong ST, Stamey TA, et al. (1994) Antiproliferative effects of 1,25-dihydroxyvitamin D3 on primary cultures of human prostatic cells. Cancer Res 54: 805–810.

139. Bikle DD, Pillai S (1993) Vitamin D, calcium, and epidermal differentiation. Endocr Rev 14: 3–19.

140. Nakagawa K, Kawaura A, Kato S, Takeda E, Okano T (2005) 1 alpha,25-Dihydroxyvitamin D(3) is a preventive factor in the metastasis of lung cancer. Carcinogenesis 26: 429–440. doi:10.1093/carcin/bgh332.

141. Chiang K-C, Yeh C-N, Chen S-C, Shen S-C, Hsu J-T, et al. (2012) MART-10, a New Generation of Vitamin D Analog, Is More Potent than 1α,25-Dihydroxyvitamin D(3) in Inhibiting Cell Proliferation and Inducing Apoptosis in ER+ MCF-7 Breast Cancer Cells. Evid Based Complement Alternat Med 2012: 310872. doi:10.1155/2012/310872.

142. Luo W, Hershberger PA, Trump DL, Johnson CS (2012) 24-Hydroxylase in cancer: Impact on vitamin D-based anticancer therapeutics. J Steroid Biochem Mol Biol. doi:10.1016/j.jsbmb.2012.09.031.

143. Omdahl JL, Bobrovnikova EA, Choe S, Dwivedi PP, May BK (2001) Overview of regulatory cytochrome P450 enzymes of the vitamin D pathway. Steroids 66: 381–389.

144. Chen G, Kim SH, King AN, Zhao L, Simpson RU, et al. (2011) CYP24A1 Is an Independent Prognostic Marker of Survival in Patients with Lung Adenocarcinoma. Clinical Cancer Research 17: 817–826. Available: http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-10-1789.

145. Norton R, O'Connell MA (2012) Vitamin D: potential in the prevention and treatment of lung cancer. Anticancer Res 32: 211–221.

146. Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. Nature 474: 609–615. doi:10.1038/nature10166.

147. Sumazin P, Yang X, Chiu H-S, Chung W-J, Iyer A, et al. (2011) An Extensive MicroRNA-Mediated Network of RNA-RNA Interactions Regulates Established Oncogenic Pathways in Glioblastoma. Cell 147: 370–381. doi:10.1016/j.cell.2011.09.041.

148. Lindholm MW, Elmén J, Fisker N, Hansen HF, Persson R, et al. (2012) PCSK9 LNA antisense oligonucleotides induce sustained reduction of LDL cholesterol in nonhuman primates. Molecular Therapy 20: 376–381. doi:10.1038/mt.2011.260.

149. Kauppinen S, Vester B, Wengel J (2005) Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. Drug Discovery Today: Technologies 2: 287–290. doi:10.1016/j.ddtec.2005.08.012.

150. Swayze EE, Siwkowski AM, Wancewicz EV, Migawa MT, Wyrzykiewicz TK, et al. (2007) Antisense oligonucleotides containing locked nucleic acid improve potency but cause significant hepatotoxicity in animals. Nucleic Acids Res 35: 687–700. doi:10.1093/nar/gkl1071.

151. Chu Y, Yue X, Younger ST, Janowski BA, Corey DR (2010) Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a

non-coding transcript at the progesterone receptor promoter. Nucleic Acids Res 38: 7736–7748. doi:10.1093/nar/gkq648.

152. Gazdar AF, Oie HK (1986) Re: Growth of cell lines and clinical specimens of human non-small cell lung cancer in a serum-free defined medium. Cancer Res 46: 6011–6012.

153. Oie HK, Russell EK, Carney DN, Gazdar AF (1996) Cell culture methods for the establishment of the NCI series of lung cancer cell lines. J Cell Biochem Suppl 24: 24–31.

154. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

155. Boutros M, Brás LP, Huber W (2006) Analysis of cell-based RNAi screens. Genome Biol 7: R66. doi:10.1186/gb-2006-7-7-R66.

156. Ding L-H, Xie Y, Park S, Xiao G, Story MD (2008) Enhanced identification and biological validation of differential gene expression via Illumina whole-genome expression arrays through the use of the model-based background correction methodology. Nucleic Acids Res 36: e58. doi:10.1093/nar/gkn234.

157. Tang H, Xiao G, Behrens C, Schiller J, Allen J, et al. (2013) A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. Clin Cancer Res 19: 1577–1586. doi:10.1158/1078-0432.CCR-12-2321.

158. Shields JM, Pruitt K, McFall A, Shaub A, Der CJ (2000) Understanding Ras: 'it ain't over 'til it's over'. Trends Cell Biol 10: 147–154.

159. Downward J (2003) Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer 3: 11–22. doi:10.1038/nrc969.

160. Corcoran DL, Pandit KV, Gordon B, Bhattacharjee A, Kaminski N, et al. (2009) Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. PLoS ONE 4: e5279. doi:10.1371/journal.pone.0005279.

161. Corcoran RB, Cheng KA, Hata AN, Faber AC, Ebi H, et al. (2013) Synthetic lethal interaction of combined BCL-XL and MEK inhibition promotes tumor regressions in KRAS mutant cancer models. Cancer Cell 23: 121–128. doi:10.1016/j.ccr.2012.11.007.

162. Scholl C, Fröhling S, Dunn IF, Schinzel AC, Barbie DA, et al. (2009) Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. Cell 137: 821–834. doi:10.1016/j.cell.2009.03.017.

163. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q, et al. (2005) KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. PLoS Med 2: e17. doi:10.1371/journal.pmed.0020017.

164. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, et al. (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second

mutation in the EGFR kinase domain. PLoS Med 2: e73. doi:10.1371/journal.pmed.0020073.

165. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, et al. (2006) GenePattern 2.0. Nat Genet 38: 500–501. doi:10.1038/ng0506-500.

166. Chin L, Andersen JN, Futreal PA (2011) Cancer genomics: from discovery science to personalized medicine. Nat Med 17: 297–303. doi:10.1038/nm.2323.

167. Chin L, Hahn WC, Getz G, Meyerson M (2011) Making sense of cancer genomic data. Genes Dev 25: 534–555. doi:10.1101/gad.2017311.

168. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19: 1639–1645. doi:10.1101/gr.092759.109.

169. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320: 1643–1647. doi:10.1126/science.1155390.

170. Miodovnik M, Koren R, Ziv E, Ravid A (2012) The inflammatory response of keratinocytes and its modulation by vitamin D: the role of MAPK signaling pathways. J Cell Physiol 227: 2175–2183. doi:10.1002/jcp.22951.

171. Haussler MR, Whitfield GK, Haussler CA, Hsieh JC, Thompson PD, et al. (1998) The nuclear vitamin D receptor: biological and molecular regulatory properties revealed. J Bone Miner Res 13: 325–349. doi:10.1359/jbmr.1998.13.3.325.