EXPLORATION OF NUCLEAR RECEPTOR ACTIVITY AND SIRNA-DERIVED PHENOTYPES AS THERAPEUTICS IN NON-SMALL CELL LUNG CANCER

APPROVED BY SUPERVISORY COMMITTEE

John D. Minna, M.D. (Mentor)

David Mangelsdorf, Ph.D. (Mentor)

Ralf Kittler, Ph.D. (Chairman)

Michael White, Ph.D.

Steven Kliewer, Ph.D.

Dedicated to:

My sweet wife Jane
My mother Helen who first taught me to love learning

EXPLORATION OF NUCLEAR RECEPTOR ACTIVITY AND SIRNA-DERIVED PHENOTYPES AS THERAPEUTICS IN NON-SMALL CELL LUNG CANCER

by

RYAN MURRAY CARSTENS

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

May, 2015

ACKNOWLEDGMENTS

I would like to thank Dr. John Minna for not throwing me out of his office when I first barraged in unannounced in the spring of 2010 and asked to rotate in his laboratory. He has provided a well-resourced and encouraging environment in which I have been able to conduct my studies and collaborate with many other world-class groups. Throughout my time in his laboratory, I have been handed more opportunities than any graduate student has a right to expect, and I cannot thank him enough for his generosity, insight, and supportive nature.

Furthermore, I would like to thank all the members of the Minna laboratory for their guidance and support. Everyone works together to make our mutual work environment a fun and inviting place, and I have greatly enjoyed being a part of the projects that have gone on here. Specifically, I'd like to than Dr. Jill Larsen for being just grumpy enough during my rotation to keep me on my guard but yet just helpful enough to not drive me away from the laboratory. Additionally, I'd like to thank her for her many technical insights and protocols early on in the project. I'd also like to thank Dr. James Sullivan for his past leadership in the laboratory and for his intentional mislabeling of all of his reagents to ward away would-be borrowers. Dr. Chunli Shao also deserves my gratitude for her friendship, timely jokes, and good natured smile. Paul Yenerall has been a willing, helpful hand with many of these experiments, and his patient forbearance of my clumsy mentoring efforts and organization of past datasets has been most appreciated. I also thank Dr. Luc Girard for his computational support and Drs. Brenda Timmons, Rebecca Britt, and Robin Frink as well as Paul Yenerall, Patrick Dospoy, and Suzie Hight for many enjoyable lunches. Finally, I also appreciate the

friendly smiles and good attitudes of the rest of the Minna laboratory--Drs. Peyton, Gao, and Das, as well as Maithili, Alex, Rachel, Dhruba, Sunny, and Sam, for their friendship and fun for the last five years.

Other UT Southwestern staff members who were particularly helpful include the office staff in the Hamon Center, Dr. Rolf Brekken for accepting me as a rotation student when I could not find my first rotation, Dr. Ralf Kittler, Dr. Steven Klewier, and Dr. Bruce Postner. Dr. Mangelsdorf who supported me as my co-mentor, and Dr. Shuguang Wei who spent countless hours making siRNA plates for me and who always goes above and beyond to ensure her work is the best possible. Also, I'd like to thank Dr. Suzanne Fuqua who has always been strongly supportive of my research efforts and an excellent collaborator.

I would especially like to thank Dr. Kenneth Huffman for many hours of delightful discussions about all things from politics to religion to the cute puppy pictures of the day to occasionally even something of a scientific nature. Buddy has been a partner with me throughout all the work of my thesis, and his constant support, ideas, and encouragement have been essential in making this work possible. I cannot thank him enough for his friendship and patience.

Outside of the UT Southwestern community, I would like to thank my sweet wife Jane for her loving support and her encouragement in my work as well as her endearing efforts to understand and be involved in my research. I furthermore recognize that none of this work would have been possible without the loving support of my parents. I thank my dad for years of patient talks and character building that enabled me to make wise decisions about both my research and everything in life, and I thank my mom for instilling in me a

deep love and appreciation for learning that led me to pursue a doctorate degree in the first place. Also, I want to thank my brothers Seth (and L'Nae), Alex (and Claire), and Zach for their love, support, and many wonderful years together growing up and many years since of laughter and fun. I also thank my community group, Brad Gessel, Braden Chapman, Andrew Laflin, Erick Fanini, Mark Greene, and Chris Slater for their support for the last three years and their guidance in my life. I know I could not have accomplished this without them or any of the other aforementioned people.

EXPLORATION OF NUCLEAR RECEPTOR ACTIVITY AND SIRNA-DERIVED PHENOTYPES AS THERAPEUTICS IN NON-SMALL CELL LUNG CANCER

Publication No. _____

RYAN MURRAY CARSTENS

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas, 2015

Supervising Professors: John D. Minna, M.D. and David Mangelsdorf, Ph.D.

Nuclear hormone receptors are master regulators of diverse cellular functions implicated in tumor pathogenesis and as oncogenic drivers of many human cancers. To better understand what role these important receptors might be playing in lung cancer, three interconnected studies were initiated to assess nuclear receptor function, expression, and drugability within the lung cancer context.

First, a "technology development" project was undertaken to produce and troubleshoot a CLIA-certifiable, high-throughput biomarker platform capable of mRNA expression signature assessment from FFPE specimens. The platform was used to assess NR/CoReg expression levels across a 500+ sample FFPE dataset. Categorical NR/CoReg downregulation upon tumor progression as well as survival benefits for patients retaining a non-pathological NR/CoReg expression pattern were discovered.

Second, a panel of 110 NR ligands was screened across a 100-member cell line panel representative of all clinically-relevant facets of lung and breast cancers to pharmacologically interrogate these receptors as novel lung cancer targets. Following completion of this screening effort, three classes of ligands targeting the estrogen, glucocorticoid, and vitamin D receptors (ER, GR, and VDR respectively) that exert anti-proliferative phenotypes on specific subsets of the lung cancer cell lines were identified. Of particular note, several of these agents are routinely used in current clinical practice (particularly dexamethasone) and represent excellent candidates for rapid clinical translation of these findings.

Finally, an RNAi-based systematic functional interrogation of NR/CoReg function was undertaken in a 100+ member cell line panel representative of all clinically-relevant facets of lung and breast cancers. A reproducible classification of lung and breast cancers was defined based on their holistic functional states as represented by the RNAi dataset. Each of these "clades" of cancer cell lines was demonstrated to be specifically targetable by unique siRNA reagents capable of inducing growth attenuation or amplification in only that clade of cell lines. Further investigation into the mechanisms of action of these siRNA reagents unexpectedly revealed that the phenotypes were largely mediated by miRNA-like seed sequence based effects rather than target-directed siRNA total complementarity silencing. Following this discovery, efforts were undertaken and subsequently completed to identify the "true" targets of these clade-specific siRNAs.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Herrington, R. A., **Carstens R.**, Omiwade O., Cavitt T.B. "Potential Biofilm Resistant Coatings." *RadTech UV/EB Technology Expo & Conf. Proceedings*, 2008.

Faulkner, R. A., **Carstens, R.**, Omiwade O., Cavitt T.B., "Biofilm Resistant Photopolymerizable Coatings." *Applied Polymer Science: Basics and Applications of Photopolymerizable Reactions.* Vol. 4 Eds. Fouassier J.P. and Allonas X. United Kingdom: Research Signpost. 2009.

Srivastava N, Kollipara RK, Singh DK, Sudderth J, Hu Z, Nguyen H, Wang S, Humphries CG, **Carstens R**, Huffman KE, DeBerardinis RJ, Kittler R "Inhibition of Cancer Cell Proliferation by PPARg is Mediated by a Metabolic Switch that Increases Reactive Oxygen Species Levels." Cell Metab. 2014 Sep

Tian H., Lu J., Shao C., Huffman K., **Carstens R.**, Larsen J., Girard L., Liu H., Canales J., Frenkel E., Wistuba I., Minna J., Hofmann S. "Systematic siRNA Screen Unmasks NSCLC Growth Dependence by Palmitoyltransferase DHHC5." *Mol. Canc. Res.* 2015

LIST OF FIGURES

Chapter 3

Chapter 4

Chapter 5

LIST OF APPENDICES

LIST OF DEFINITIONS

Androgen Receptor	AR
Breast Cancer	BRC
Co-Regulator	CoReg
Carbamoyl Phosphate Synthetase I	CPS1
Dexamethasone	DEX
Eukaryotic Translation Initiation Factor A3	eIF4A3
Estrogen Receptor α or β	ERα, ERβ
Formalin Fixed Paraffin Embedded	FFPE
Glucocorticoid Receptor	GR
Hormone Response Element	HRE
microRNA	miRNA
Metastasis Associated 1	MTA1
(3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium)	MTS
Nuclear Receptor	NR
Nuclear Receptor Related Protein 1	NURR1
Principle Component Analysis	PCA
Poly(RC) Binding Protein 1	PCBP1
Protein Arginine Methyltransferase 1	PRMT1
Quantitative Nuclease Protection Assay	qNPA

RPMI-1640 + 5% Fetal Bovine Serum (FBS)	R5
Small Interfering RNA	siRNA
Selective Nuclear Receptor Modulators	SNuRMs
Steroid Receptor Co-Activator 3	SRC-3

CHAPTER ONE: Literature Review of Relevant Background Topics

1.1 Nuclear Receptors: Master Regulators of Cellular Function

Nuclear Receptors (NRs) are a unique class of proteins which serve as global regulators of many cellular functions. There are 48 members of the NR superfamily which conventional phylogenic studies divide into seven subgroups (1999). An alternate method of classifying NRs, cladistic analysis of anatomical profiling of NR transcript levels, finds six, similar subgroups--suggesting the conserved structure of these receptors plays a strong role in their function (Bookout, Jeong et al. 2006).

Evolutionary, NRs are found only in metazoans with different species exhibiting a wide range in terms of the number of NR family members they possess (for example, *C. Elegans* has 284 NRs while rats have 47, mice have 49, and humans have 48) (Escriva, Langlois et al. 1998; Zhang, Burch et al. 2004). NRs contain five structural domains: (A) an N-terminal domain, (B) a DNA binding domain, (C) a hinge region, (D) a ligand binding domain, and (E) a C-terminal domain (Kumar and Thompson 1999). The presence of a DNA binding domain allows NRs to function as transcription factors, bind DNA, and directly regulate the expression of adjacent genes (Tang, Chen et al. 2011).

1.1.2 The Unique Mechanism of Nuclear Receptors

One of the key differences between NRs and other transcription factors is the NR's ligand binding domain. Typically, the binding of a ligand (a lipophilic substance such as a hormone, or a xenobiotic endocrine disruptor) induces an allosteric change in the conformation of the receptor which activates it, leading to the up and down regulation of genes (Schupp and Lazar 2010). Ligands are typically classified according to their mechanism of action (broadly, agonist or antagonist), the receptor(s) they bind, and whether or not they are naturally occurring (Sladek 2011).

NRs fall into four broad classes in terms of their mechanisms of action in response to ligands (Mangelsdorf, Thummel et al. 1995). Type I receptors reside as monomers bound by complexes in the cytosol until a ligand binding event leads to homo-dimerization of the receptor and translocation to the nucleus where it binds specific hormone response elements (HREs) and subsequently exerts its effects. In contrast, Type II receptors exist in the nucleus regardless of their ligand binding status and exert their effects as hetero-dimers. The activity of the receptor changes in response to ligand binding typically through the exchange of repressive Co-Regulator (CoReg) proteins for activating CoReg proteins. Type III and IV receptors are similar to Type I differing mainly in the types of HREs they bind (an inverted repeat for Type I, a direct repeat for Type III, and a binding of only a single half site for Type IV).

Because of their unique mechanism, NRs represent the idealized drug target, a protein which naturally possesses a lipophilic small molecule binding pocket and changes its activity in response to a binding event(Aranda and Pascual 2001). Given this, it is no surprise that the NR superfamily is the most targeted family of proteins (drugs/per family member) with 13% of all FDA approved drugs targeting one of these receptors (Overington, Al-Lazikani et al. 2006). NR targeting ligands can be endogenous, xenobiotic, or synthetic in origin and, although differing definitions of "ligand" makes concrete numbers difficult, it can be safely stated that more than 200 bona fide NR ligands have been identified to date (Sladek 2011), (www.nursa.org).

1.1.3Various Mechanisms of Ligand Action

Ligands can exert their effects through a variety of mechanisms each of which has a specific nomenclature (Germain, Staels et al. 2006).

(A) Agonists: Ligands classified as agonists simply bind the receptor and lead to its activation (typically through locking the receptor in the active conformation). Some examples of NR agonists include estradiol's action on estrogen receptors α and β (ERα, ERβ) and dihydrotestosterone on androgen receptor (AR).

(B) Antagonists: Loosely, antagonists are compounds which prevent a nuclear receptor from adopting an active conformation. This can occur through a variety of mechanisms including bulky side chains that prevent co-activator binding (AF-2 antagonists, exemplified by BMS614 and Retinoic Acid Receptor alpha)

and inactive conformations of the receptor induced by the antagonist binding event (such as THC's mechanism of action on ERβ) (Germain, Iyer et al. 2002; Shiau, Barstad et al. 2002).

(C) Inverse Agonists: Inverse Agonists are synthetic ligands capable of preventing the basal activity of NRs even in the absence of exogenous agonist. This can be accomplished through the recruitment of Co-Repressors or though induced conformational changes (Greschik, Flaig et al. 2004).

(D) Selective Nuclear Receptor Modulators: Selective Nuclear Receptor Modulators (SNuRMs) represent an class of ligands that exhibit either agonist or antagonist activity dependent on the environmental context. Because of their unique mechanism, SNuRMs have proven invaluable when attempting to overcome side effects associated with NR targeted drugs. Tamoxifen is the prototypical example of a SNuRM which acts as an antagonist in breast tissue and an agonist in uterus due to higher SRC-1 levels in uterine tissue (Shang and Brown 2002). Typically, the differential expression of CoRegs is responsible for the specificity of SNuRMs (Smith and O'Malley 2004).

1.2Co-Regulators: Partners with Nuclear Receptors

Co-Regulators (CoRegs) are a much larger family of proteins than nuclear receptors (NRs) consisting of more than 300 members (Lee, Lee et al. 2001). Unlike nuclear receptors, the family is not defined by structural similarity, but rather by functional similarity.

Specifically, a CoReg is any molecule that interacts with a nuclear receptor for the purpose of either (1) enhancing the NR's transactivation (termed a coactivator) or (2) lowering the transcription rate at the NR's target genes (termed a corepressor) (McKenna, Lanz et al. 1999). Initially, each CoReg was typecast as either a coactivator or a corepressor, but more recent studies have blurred the lines between the distinctions to some degree (O'Malley and McKenna 2008).

CoRegs are an answer to how nuclear receptors can simultaneously control the activation and repression of hundreds of genes within a given cell (Glass and Rosenfeld 2000). Furthermore, the differential expression of CoRegs in different tissue types allows for the same nuclear receptor to have completely opposing effects on the same gene in different contexts (Aranda and Pascual 2001).

1.2.1 General Model of CoRegulator Function

Much variation exists among CoRegs as to the methods by which they are regulated (such as phosphorylation (Rowan, Weigel et al. 2000) or ubiquitination (Lonard, Nawaz et al. 2000)) and the mechanisms by which they influence NR action. However as a general rule, CoRegs exist in large complexes in a cell with these complexes then interacting directly (protein-protein) with NRs (McKenna, Nawaz et al. 1998). CoRegs can receive signals from a large number of cellular signaling pathways including cross-talk from G protein coupled receptors and tyrosine signaling pathways (McKenna and O'Malley 2002).

1.2.2 CoRegulators and Cancer

Despite their relatively recent identification, important roles for many CoRegs (both coactivators and corepressors) have been identified within the cancer context (O'Malley and Kumar 2009). Specifically, more than half of the identified CoRegs to date have been found to have roles in various diseases (presumably because of their regulatory nature on NRs) with the most common being the cancer context (O'Malley 2006). Some of the most highly studied CoRegs within the cancer context include SRC3, MTA1, and PCBP1. SRC3 is amplified in approximately 10% of breast cancers and overexpressed in more than half of all cases (Anzick, Kononen et al. 1997). Mouse knockout/overexpression studies with SRC3 have confirmed its role as an oncogene within the cancer context (Torres-Arzayus, Font de Mora et al. 2004). Similar to SRC3, MTA1 has also been found to be overexpressed in a variety of tumors (Manavathi and Kumar 2007). MTA1 has been shown to work in concert with estrogen receptor alpha (ERa) to allow the estrogen-independent growth of tumor cells (Manavathi and Kumar 2007).

1.3 Cancer

Given its prevalence and impact on modern society, cancer as a disease scarcely needs introduction. In 2012, more than fourteen million new cases of cancer occurred globally and cancer was responsible for more than eight million deaths worldwide (approximately 14% of all human deaths) (*World Cancer Report,* WHO 2014). Cancer is also one of the oldest known diseases with tumors having been discovered in Egyptian

mummies (David and Zimmerman 2010). It has been described continuously throughout all medical history with reports being found in the writings of many diverse cultures from antiquity until modern day (Hajdu 2011). Furthermore, the burden of cancer is increasing in virtually all developing/developed countries – presumably due to changes in diets, physical activity levels, air quality, exposure to carcinogens, and generally increased life expectancies (Jemal, Bray et al. 2011).

1.3.1 Characteristics of Human Cancer

Broadly defined, cancer is considered the abnormal growth of cells (caused by multiple changes in gene expression leading to a dysregulated balance of cell proliferation and cell death) that ultimately evolves into a population of cells that can invade tissues and metastasize to distant sites, causing significant morbidity and, if untreated, death of the host. Our understanding of the disease has changed significantly in the last twenty-five years with cancer now being largely considered a disease driven by genetic changes (Tamborero, Gonzalez-Perez et al. 2013).

In reality, "cancer" is an umbrella term used to refer to many subsets of disease. Each sub-disease can be distinguished by variations in age of onset, rate of growth, state of cellular differentiation, invasiveness, response to treatment, prognosis, and (arguably most importantly) tissue of origin (Greystoke and Mullamitha 2012). Further sub-divisions exist within each tissue site which can be defined histopathologically or by the increasingly more

common molecularly defined methods as are commonly used in breast cancer (ER/PR/HER2 status) or in lung cancer (KRAS/EGFR mutation status) (Engstrom, Opdahl et al. 2013), (Benesova, Minarik et al. 2010).

Mortality rates vary greatly between different cancer types with some having a good prognosis (such as breast (~80% survive) or prostate (~85% survive) cancers) and others being particularly deadly (such as lung (~5-10% survive) or colon (~10% survive) cancers) (Siegel, Naishadham et al. 2013). Furthermore, mortality rates for different cancer subtypes have changed dramatically over the last 100 years with deaths due to stomach cancer today being half of what they were in the 1930s (presumably due to the introduction of preservatives) and lung cancer deaths having increased by five to ten fold during the same time period (largely due to increased smoking rates) (Ngoan, Mizoue et al. 2002).

1.3.2 Causes of Cancer

Surprisingly despite centuries of study, researchers still do not have a comprehensive picture of the cause(s) of cancer. However, some general concepts have been understood that have been applied to great benefit. Specifically, it is know that genetic alterations (mutations, copy number changes, deletions, rearrangements) are responsible for much (but not all) of human cancer with two genetic events generally being considered the minimum number for early stage cancers to develop (Knudson 2001). An extensive list has now been delineated of genes known to contribute either positively (oncogenes) or negatively (tumor suppressors) to the genesis of cancer with new genes being added regularly (Lengauer,

Kinzler et al. 1998). Genetic events can occur through a variety of mechanisms including heredity, viral infections, metabolic-associated oxidative stress, radiation induced DNA brakes, and chemical carcinogens (Jackson and Loeb 2001). Our understanding of carcinogens has led to restrictions or bans on the use of many chemicals (Reviewed: (Siemiatycki, Richardson et al. 2004)) with the most famous case being the linking of cigarette smoking to lung cancer (Sasco, Secretan et al. 2004).

1.3.3 The Epidemiology of Human Cancer

Epidemiological studies of cancer have been invaluable in furthering our understanding of cancer biology. Some key epidemiological findings will be reviewed.

The most basic epidemiological finding surrounding cancer is the association of cancer incidence with age (average age for all sites is 67) (Gloeckler Ries, Reichman et al. 2003) with the large majority of cancers occurring either early (developmental/hereditary cancers) or late in life (induced cancers due to genetic events). Accounting for the most cancer deaths in men, the most common cancer worldwide is lung cancer--which tracks closely with smoking incidence with a lag time of approximately 20 years in both men and women (Loeb, Ernster et al. 1984).

Also of importance, the prevalence of different cancer types varies significantly across regions with some notable examples being the disproportionately high rates of bladder cancer in Africa, liver cancer in China, and colorectal cancer in the developed world (Kamangar,

Dores et al. 2006). Furthermore, the molecular subtypes found within a given disease can vary greatly by region. For example, EGFR mutations are found much more commonly in Asian patients than in patients of European or North American origin (Boch, Kollmeier et al. 2013).

Within each cancer, epidemiological studies have also revealed that the underlying driving factors vary greatly between different cancer types. For example, KRAS mutations are common in lung, liver, and colon cancers while breast cancer is often driven by hormonal factors or loss of tumor suppressors such as BRCA1 (Wang, Kaiser et al. 2013), (Carter 2001). In contrast, cervical cancer is most commonly associated with sexually transmitted infection with certain species of human papilloma virus (Crosbie, Einstein et al. 2013).

1.3.4 The Biochemistry and Cell Biology of Cancer

Since the 1920s work of Otto Warburg, it has been appreciated that there are differences between the biochemistry of cancerous cells and normal ones (Warburg, 1930). Subsequent studies have further defined the common phenotypes of cancerous cells and broadly characterized them into ten categories known as the hallmarks of cancer (Hanahan and Weinberg 2011).

First and foremost, cancer cells must be able to continuously divide – a property referred to as "replicative immortality." When cells divide many times, they can encounter

problems due to telomere (the end caps of chromosomes) shortening. The majority of cancer cells avoid this problem by activating telomerase, an enzyme capable of maintaining telomere integrity, though other possible mechanisms have been reported (Shay and Wright 2011). Because telomerase is normally expressed only in rare stem cells, this property of cancer has led some to hypothesize that stem cells may be preferentially selected as the cell or origin for many cancers (Sell 2004; Bu and Cao 2012).

Along similar lines, cancer cells must maintain a fine balance between mutation-inducing genome instability and prevention of the cell's ability to detect the DNA damage it has incurred. Because cancers are generally caused by DNA damage, breakage of the cell's DNA repair machinery is almost universal among cancers. Of particular note, recent sequencing studies have identified a new subclass of cancers (characterized by an ultra-high number of mutations) driven exclusively by a pro-mutation environment permitted through mutations in POLE (Palles, Cazier et al. 2013).

The mass DNA damage incurred by most cancers would typically not be permitted in a cell except that cancers universally develop an ability to sabotage or circumvent the cell's attempts to apoptose (commit cellular suicide). Resisting cell death is a key hallmark of cancer and can be accomplished through many different (often overlapping) mechanisms. The most common apoptosis avoidance mechanism is the inactivation of p53 which is observed in more than 50% of human cancers (Hollstein, Sidransky et al. 1991). Other

mechanisms include mutation or deletion of key genes in the apoptotic machinery or upregulation of negative regulators of apoptosis (such as BCL-2) (Ashkenazi 2002).

In addition to preventing apoptosis signals, a cancer cell must also prevent growth suppressive signals. Typically, these are encountered in the form of cell cycle checkpoints or negative regulators of the cell cycle (Rabbani and Cordon-Cardo 2000). Cell cycle machinery has been well-defined through countless studies (Murray 2004) and mutations can be found scattered throughout nearly all the "brakes" on this system with notable recurrent mutations occurring in Rb, CDK4, CDK6, CCND2, and CDKN2 (Nojima 1997).

While disconnecting the "brakes" on the cellular growth machinery is important, equally important for a cancer cell is to obtain a sustaining proliferative signal. Frequently referred to as driving mutations, these proliferative signals receive particular attention because much research has shown that cancers can often become "addicted" to these growth signals and that removal of the signal can lead to the rapid death of the cell (Weinstein and Joe 2006). There are many sources of proliferative signals that vary among different cancer types that can be arranged into common pathways (Jones, Zhang et al. 2008). Some examples include KRAS or EGFR mutations (Benesova, Minarik et al. 2010), PIK3CA/PTEN alterations (Gonzalez-Angulo, Ferrer-Lozano et al. 2011), androgen receptor signaling in prostate cancer (Heinlein and Chang 2004), and estrogen receptor signaling in breast cancer (Sommer and Fuqua 2001).

As the pro-proliferative signals take effect, cancer cells quickly encounter atypical energy requirements for their growth and must change their metabolism accordingly – altering their production and consumption rates and methods of synthesis of carbohydrates, proteins, lipids, and nucleic acids. Furthermore, cancer cells shift from oxidative phosphorylation to glycolysis for ATP production (the Warberg effect and the basis for PET imaging of tumors) presumably to reduce oxidative stress and to allow for more growth under the hypoxic conditions found in many tumors (Annibaldi and Widmann 2010). The first recognized change in cancer cells, deregulation of cellular energetics, remains an active topic of research, an active target for new therapeutics, and a mysterious segment of cancer biology that is not yet fully characterized (Cairns, Harris et al. 2011).

As a tumor continues to grow, it encounters new problems resulting from its increased resource consumption. First among these new problems is the need for an increased blood supply (Otrock, Mahfouz et al. 2007). As tumors grow larger, the diffusion limitations of oxygen and nutrients become too great, and cells in the center of the tumor become necrotic and do not receive enough oxygen to sustain their growth (Carmeliet 2005). To combat this problem, tumors initiate a process known as angiogenesis – the growth of new blood vessels (Tahergorabi and Khazaei 2012). Several agents have been developed targeting this process most notably among which is Bevacizumab, an antibody that works by inhibiting Vascular Endothelial Growth Factor-A (VEGF-A) (Shih and Lindley 2006).

Eventually, even the local resources available following angiogenic processes are insufficient to allow for the continued growth of the tumor, and the cells begin the process of invading to new sites. This process, known as metastasis, is what is ultimately responsible for most cancer morbidity and mortality as these new growths can often occur in more perilous locations such as the brain or liver (Valastyan and Weinberg 2011). Although models vary, the typical understanding of metastasis is referred to as epithelial-to-mesenchymal transition (EMT) whereby cells alter their properties to allow them to survive outside of their niche and away from adhesion signals (Leber and Efferth 2009). It is also commonly hypothesized that cells which have undergone EMT are more "stem-like" and have increased drug resistance properties (Singh and Settleman 2010).

Finally, underlying every step of the oncogenic process is the requirement for a cancer cell to evade immune detection. Although it is difficult to study, the increased incidence rates of cancer in immunosuppressed/immunocompromised individuals strongly suggests that the immune system actively detects and eliminates early stage cancers (Grulich, van Leeuwen et al. 2007). Furthermore, tumor associated inflammation caused by immune cell invasion/infiltration into tumors plays a key role in tumor cell proliferation and immune tolerance (Philip, Rowley et al. 2004). All of these concepts have recently received much attention following the successful clinical trial results with anti-PD-1 antibodies (Nivolumab)

which are the first clinical demonstration that breaking immune tolerance can lead to tumor regressions (Topalian, Hodi et al. 2012).

1.4 Lung Cancer

1.4.1 Statistics

Worldwide, lung cancer is the most common cause of cancer-related death in men and women accounting for more than 1.5 million deaths in 2012 – more than breast, colon, prostate, and pancreatic cancers combined (Siegel, Naishadham et al. 2013). Considering all stages together, lung cancer has an approximately sixteen percent five-year survival rate with early stage tumors generally having good prognosis (>80% survival rates) and late stage tumors being nearly universally fatal (Henschke, Yankelevitz et al. 2006).

The largest risk factor for developing lung cancer is undisputedly smoking with more than eighty percent of lung cancer cases occurring in patients reporting long-term exposure to tobacco smoke (Boyle 1997). Risk of lung cancer development is approximately eleven to twenty-two times higher for smokers. A dose-response relationship exists whereby heavy smokers are more likely than light smokers to develop lung cancer, and cessation of smoking results in lowered risk of mortality from lung cancer versus patients who do cease smoking (Shopland, Eyre et al. 1991). Also noteworthy, lung cancer pathophysiology in never smokers has been found to be significantly different than in smokers – harboring fewer

mutations overall, higher incidence of EGFR mutations, and generally different pathologies (Sun, Schiller et al. 2007).

1.4.2 Molecular Epidemiology

Not surprisingly, "lung cancer" is an umbrella term encompassing many diverse disease states. The most common subdivision is accomplished through pathological analysis and breaks lung cancer down into either small-cell lung carcinoma (SCLC, roughly twenty percent of cases) or non-small-cell lung carcinoma (NSCLC, roughly eighty percent of cases) (Jemal, Bray et al. 2011). NSCLC can be further broken down into three subgroups: (1) adenocarcinoma - approximately forty percent of lung cancers, most common subtype among people classified as "never-smokers," (2) squamous-cell carcinoma - approximately thirty percent of lung cancers, and (3) large-cell carcinoma - approximately ten percent of lung cancers (Travis 2011). Other subtypes exist that are not generally officially recognized (such as neuroendocrine) but which may have important clinical implications (Varlotto, Medford-Davis et al. 2011).

Increasingly commonly, lung cancer is also sub-classified by the particular molecular markers found in a given instance regardless of histotype. The most well-studied of these is EGFR which drives ten to thirty percent of lung cancers (depending on which patient population is studied--EGFR mutations are much more common in Asian populations) and is the molecular marker defining EGFR inhibitor sensitivity (da Cunha Santos, Shepherd et al. 2011). Equally important, KRAS mutations are found in approximately twenty to thirty

percent of lung cancer patients and represent a particularly challenging form of the disease as there is currently no targeted therapy for these patients (Roberts and Stinchcombe 2013). Other identified drivers include HER2, BRAF, PIK3CA, AKT1, MAP2K1, and MET with each making up a small percentage (less than five percent) of patients (Pao and Girard 2011). Trials are underway testing targeted therapies for each of these subsets.

1.4.3 Treatment

The most common treatments for lung cancer include palliative care, surgery, chemotherapy, radiation therapy, and (increasingly) targeted therapies (Cagle and Chirieac 2012). NSCLC and SCLC are typically treated differently with surgery being quite common in NSCLC while SCLC receives only chemo/radiation combinations (Cooper and Spiro 2006). Chemotherapy regiments for lung cancer almost universally include a platinum-based agent combined with either etoposide for SCLC or a taxane for NSCLC. Generally, radiotherapy is also included – especially when surgical options are unavailable (Goffin, Lacchetti et al. 2010) .

In cases where oncogene addiction can be demonstrated by presence of certain molecular markers or fusion products, targeted therapies have proven very successful (Larsen, Cascone et al. 2011). Unfortunately, targeted therapy treatment is almost universally accompanied with relapse and resultant recurrent tumors are much more refractory to further treatment (Tang, Salama et al. 2013).

1.4.4 Nuclear Receptors and Lung Cancer

As has been discussed, lung cancer is a strikingly heterogeneous disease and, as such, comprehensive characterization of its pathobiology has proven elusive. Of particular note, little work has been done to characterize a comprehensive role for nuclear receptors in lung cancer. Scattered reports exist in the literature of various NR ligands potentially having effects in lung cancer (particularly retinoids (example: (Fritz, Kennedy et al. 2011)) and glucocorticoids (example: (Greenberg, Hu et al. 2002))), but none have been conclusive.

Preliminary work by the Minna and Mangelsdorf laboratories has sought to offer a broad scale view of NRs in lung cancer. Initial studies have profiled the expression levels of the forty eight nuclear receptors across a panel of lung cancer cell lines and demonstrated 1) NR expression patterns are different between tumor and normal samples, 2) NR expression patters can subtype lung cancer specimen, and 3) NR ligands can be effective in some cases at eliciting an antitumor response (Jeong, Xie et al. 2012). Further studies demonstrated the potential of NR expression signatures to provide prognostic information (Jeong, Xie et al. 2010).

1.5 RNAi Technology

In 1998, it was discovered in *C. Elegans* that that mRNA could be specifically targeted by carefully designed, double stranded RNA molecules resulting in target cleavage, a discovery which led to a quick Nobel Prize award just eight years later (Fire, Xu et al.

1998). siRNAs were then discovered a year later as a natural phenomenon in plants (Hamilton and Baulcombe 1999) and synthetic siRNAs were first used to silence specific target transcripts in mammalian cells two years later (Elbashir, Harborth et al. 2001). Since then, the transformative power of this technology has been quickly realized, and genome-wide libraries have been developed and used to screen the functional relationships of mRNAs to a variety of phenotypes, particularly focusing on the viral infection and oncology fields where functional relationships are key (Cullen 2006).

Extensive research has gone into optimizing the structure of synthetic siRNAs with short (usually 21 base pair) double-stranded RNAs featuring 5' phosphorylated ends and hydroxylated 3' ends with two overhanging nucleotides proving the most effective design. This design compromises between being too short for effective target specificity and too long such that cells activate an interferon response (Sioud 2006; Olejniczak, Galka et al. 2010). Naturally produced siRNAs are processed by the dicer enzyme from long double-stranded RNAs and short hairpin RNAs while synthetic siRNAs are transfected into cells through a lipid-based transfection reagent protocol (Bernstein, Caudy et al. 2001). RNAi specificity is excellent with even single-nucleotide base changes completely abolishing activity (Du, Thonberg et al. 2005).

With more than twenty-nine genome-wide siRNA screens having been completed by 2011 (Sigoillot and King 2011), there can be no doubt that siRNA technology has had a

major impact on our understanding of functional biology. Of particular note, three screens were performed in the context of HIV (Brass, Dykxhoorn et al. 2008; Zhou, Xu et al. 2008; Yeung, Houzet et al. 2009) to identify host factors targetable for the treatment of the disease (meta-analysis of the three screens (Bushman, Malani et al. 2009)). Similarly, multiple independent screens were done to identify targets exhibiting synthetic lethality in conjunction with KRAS mutation (Sarthy, Morgan-Lappe et al. 2007; Barbie, Tamayo et al. 2009; Luo, Emanuele et al. 2009; Scholl, Frohling et al. 2009).

1.5.1 RNAi and the miRNA Machinery: Initial Discoveries

As early as 2003, just two years after the first report of a synthetic siRNA in mammalian cells, it was discovered that siRNAs were capable of using more of the miRNA machinery than just the dicer enzyme and that they were, in fact, capable of functioning as miRNAs (Doench, Petersen et al. 2003). In a landmark paper, it was also reported in 2003 that siRNAs could downregulate other targets in addition to their intended target (Jackson, Bartz et al. 2003). Months later, a second group independently confirmed the finding (Semizarov, Frost et al. 2003), and studies were undertaken to further refine the technology to overcome these "off-target" silencing events and to elucidate the mechanisms that might control them.

Following these initial reports, much was learned about the interactions of synthetic siRNAs with the miRNA machinery. Work in 2005 discovered that siRNA-mediated off-target silencing was triggered by a 7-nucleotide portion of the siRNA (later termed the

siRNA's "seed sequence") (Lin, Ruan et al. 2005), and a second study identified Argonaute2 as the protein responsible for cleavage of the guide strand of the siRNA during RISC activation, explaining why only one of the potential two seed sequences was causing off-target silencing through the miRNA machinery (Rand, Petersen et al. 2005).

1.5.2 RNAi miRNA-Like Off-Target Effects: A Wide-Spread Phenomenon

In 2006, the floodgates were opened and many papers were published including several highly influential reviews (Echeverri and Perrimon 2006; Ma, Creanga et al. 2006) and methods papers (Cullen 2006; Echeverri, Beachy et al. 2006) detailing proper techniques to ensure the accurate interpretation of RNAi data. It was discovered that off-target effects were highly prevalent in Drosophila RNAi screens as well as those done in mammalian cells (Ma, Creanga et al. 2006), and that off-target effects were due to 3' UTR seed matches, not the overall identity of the siRNA to the off-target transcript (Birmingham, Anderson et al. 2006; Jackson, Burchard et al. 2006).

Since 2006, consensus has been reached that miRNA-like off-target effects pose a real problem for siRNA screens. Much work has been done on elucidating the exact mechanisms by which seed sequences (Anderson, Birmingham et al. 2008; Brodersen and Voinnet 2009) and other parts of the siRNA (Grimson, Farh et al. 2007; Dahlgren, Zhang et al. 2008) contribute to off-target selection. Other studies have focused on better defining the scale of the problem with array profiles being assessed for the non-targeting control siRNA against

GFP (Tschuch, Schulz et al. 2008) or the mechanisms by which the miRNA machinery selects its targets (Czech and Hannon 2011).

1.6 Strategies to Overcome RNAi Based Off-Target Effects

Since as early as 2005-almost immediately following the first reports of off-target effects-much research has been conducted in an effort to produce new reagents or strategies to overcome the challenges presented by siRNAs silencing transcripts in addition to their intended targets. These efforts can be broadly categorized as (1) novel siRNA chemistries/structures, (2) alternative RNAi triggers, (3) better experimental designs, and (4) algorithms to predict and avoid off-target events. Each will be reviewed below.

1.6.1 Novel siRNA Chemistries/Structures.

Initial work in the field of novel siRNA chemistries/structures focused on the use of locked nucleic acid (LNA) technology to improve both the stability of the siRNA (to aid in *in vivo* delivery) and to reduce off-target events (Elmen, Thonberg et al. 2005). LNA technology modifies the ribose moiety with an extra bridge connecting the 2' oxygen and the 4' carbon effectively "locking" the ribose and reducing flexibility/increasing the melting temperature of the oligo (Kaur, Arora et al. 2006; Owczarzy, You et al. 2011). Although this technique seemed initially promising, further study demonstrated reduced silencing efficacy of these oligos, and difficulties in commercialization due to restrictive licensing options with patents prevented widespread use of LNAs in the siRNA field (Mook, Baas et al. 2007).

Structural analysis of siRNAs discovered that the 2'-OH was not required for siRNA activity (Chiu and Rana 2003). This finding has led to many different modifications of siRNAs at this position for a variety of purposes (Reviewed: (Watts, Deleavey et al. 2008)), the most common and frequently utilized being 2'-O-Methylation (2'-O-Me). 2'-O-Me is now the basis of Dharmacon's "On-Target Plus®" product technology (one of the most popular siRNA products currently available) following a 2006 report demonstrating reduced numbers of downregulated "off-target" transcripts when using 2'-O-Me modified siRNAs (Jackson, Burchard et al. 2006). Subsequent studies have been conflicting as to the efficacy of 2'-O-Me siRNAs with most groups reporting reduced silencing efficiency (Elbashir, Martinez et al. 2001; Braasch, Jensen et al. 2003; Chiu and Rana 2003; Czauderna, Fechtner et al. 2003) although other reports see little loss of activity (Choung, Kim et al. 2006). Presumably, these differences in efficacy are due to differences in experimental design (Kraynack and Baker 2006).

To summarize, it is accepted in current literature that chemical modifications of siRNAs are capable of reducing off-target profiles and improving *in vivo* delivery mechanics of siRNAs. However, it is also clear that no current modification scheme has been able to effectively *solve* either the off-target issue or *in vivo* immunogenicity concerns (Reviewed: (Engels 2013)).

1.6.2 Alternative RNAi Triggers

Since their first identification in 2003 (Jackson, Bartz et al. 2003), off-target effects have led researchers to investigate alternative methods for triggering RNA-interference events that might achieve the same results without propagating unintended silencing events. The two main classes of reagents studied under this field are short-hairpin RNAs (shRNA) and endoribonuclease prepared siRNAs (esiRNA).

Short hairpin RNAs (shRNAs) are a sequence of RNA that enters the cellular RNAi machinery and achieves gene silencing (Reviewed: (Rao, Vorhies et al. 2009)). Unlike siRNAs, shRNAs are delivered through an expression vector which can be transfected into the cell as a plasmid or, more commonly, transduced using a viral or bacterial vector system (Singer and Verma 2008). This vector then drives expression of the shRNA which is then further processed by cellular machinery to ultimately yield siRNA (Tijsterman and Plasterk 2004). shRNAs were first developed in 2002 (almost immediately after the initial reports that siRNAs worked in mammalian systems) as a method to allow for more inexpensive production of siRNAs as well as to allow for *in vivo* studies using RNAi (Brummelkamp, Bernards et al. 2002; Paddison, Caudy et al. 2002).

Despite their widespread use, far fewer studies have been conducted to assess the potential that shRNAs might have for causing off-target effects than have been conducted with siRNAs (Reviewed: (Rao, Senzer et al. 2009)). However, one key study conducted in 2010 reported reduced off-target silencing from shRNAs although a robust off-target

signature could still be detected with the reagent (Klinghoffer, Magnus et al. 2010). One possible mechanism explaining this reduction in unintended silencing events was given with the 2012 discovery that shRNA processing by Dicer is not as robust as previously thought and multiple different siRNAs could be produced from the same shRNA under certain conditions (Gu, Jin et al. 2012). Another safeguard employed by shRNA users is inherent to the typical shRNA screen design whereby many shRNAs targeting the same transcript are used independently and then the results are viewed holistically to consider if each independent reagent is producing the same phenotype (suggestive of an on-target effect).

First introduced in 2002, endoribonuclease prepared siRNAs (esiRNA) offer a unique approach to solving the off-target effects issue (Yang, Buchholz et al. 2002). esiRNA reagents are produced by the targeted digestion – using RNAse III - of an amplified target sequence (typically 400 to 600 base pairs in length). Because miRNA specificity can be altered by a single frame-shift in the seed region, esiRNA pools represent approximately four to six hundred independent siRNAs, each with their own seed sequence, that all target the original transcript (Kittler, Surendranath et al. 2007). The technology has been used successfully on the genome scale (Kittler, Pelletier et al. 2007; Ding, Paszkowski-Rogacz et al. 2009), in Zebrafish (Liu, Wang et al. 2005), and in more targeted experiments (Kolas, Chapman et al. 2007; Lawo, Bashkurov et al. 2009).

Although esiRNA reagents do convincingly reduce seed-based effects almost to the point of elimination (Kittler, Surendranath et al. 2007), these reagents unfortunately come with their own set of technical challenges. Specifically, the enzymatic digestion step with RNAse III produces fragments of various sizes--the largest of which are capable of triggering an interferon response in cells. The interferon response is a highly-studied phenomenon whereby cells detect double stranded RNA (dsRNA) using complex sensors and undergo self-destruction in response, presumably to help prevent viral propagation (Reviewed: (Olejniczak, Galka et al. 2010)). Because of their specific 21 nucleotide size, siRNAs are typically able to avoid triggering this response. However, the 25 nucleotide fragments sometimes resultant following esiRNA production are often sufficient to trigger interferon responses, especially in particularly susceptible tissues. Because of this drawback, esiRNA has seen limited use and it remains to be seen if further technical advancement will be achieved with esiRNA technology to allow for its widespread application.

1.6.3 Better Experimental Designs

One of the simplest and most common methods employed to avoid off-target effects involves basic changes to experimental protocols which are thought to reduce, avoid, prevent, or detect phenotypes resultant from off-target silencing events. The first strategy developed (which is today almost universally employed) is the pooling of siRNA reagents targeting the same transcript (Echeverri and Perrimon 2006). Similar to the concept behind endoribonuclease prepared siRNAs (esiRNAs), pooling synthetic siRNA reagents aims to

reduce off-target events by introducing a variety of seed sequences while continuing to maintain a single on-target transcript (Myers, Chi et al. 2006). Although it is commonly acknowledged that using pools with fewer than 10 or 20 independent siRNAs is unable to fully prevent off-target silencing, the practice of performing pooled siRNA screens still continues. Some reasons that have been cited for the continued use of pooled siRNAs in screening protocols include that using pools might help reduce the magnitude of *some* off-target events, and that using a pool of four siRNAs may increase knockdown efficiency of the intended target. Interestingly, it has now been experimentally shown that screens which use pools of siRNAs find more phenotypic results than those using individual siRNAs (Parsons, Schindler et al. 2009).

The other common experimental design technique employed by researchers for enhancing confidence in the specificity of phenotypic results is the validation of results by a second silencing reagent targeting the same transcript. Used as a common standard for publication, "validation by a secondary silencing trigger" has been recommended as sufficient proof of on-target activity in lieu of the gold standard genetic rescue experiment (Cullen 2006; Mohr, Bakal et al. 2010). While this method adds some confidence in the on-target nature of the reported results, it has been demonstrated to be far from sufficient in many cases (Sigoillot and King 2011).

1.6.4 Algorithms to Predict and Avoid Off-Target Effects

A different approach to avoiding miRNA-based off-target effects is the implementation of bioinformatics tools to try and design siRNA sequences that are less likely to cause off-target events (reviewed: (Tilesi, Fradiani et al. 2009)). While many of papers touting tools along this vein have been published, each approach can be broadly categorized as follows:

(A) Functional Off-Target Filtering: Functional off-target filtering is a two-step methodology that tries to use known relationships between proteins to predict off-target events. First, the algorithm identifies a list of transcripts that could potentially be downregulated by a given siRNA through a miRNA-like mechanism. This list is then cross-compared to previously-derived lists of genes that are thought to have the same functional mechanism as the on-target of the siRNA being designed (Das, Ghosal et al. 2013). *(To clarify the concept with an example: In a siRNA screen for novel regulators of the TGFb pathway, test siRNAs which are predicted through computational methods to target known regulators of the TGFb pathway would be thrown out of the analysis as they would be suspected to be causing their phenotypes through an off-target mechanism.)* Functional filtering has been used in many other off-target algorithms (examples: (Iyer, Deutsch et al. 2007; Wang, Varma et al. 2009; Mysara, Garibaldi et al. 2011)) but has the obvious limitation of being ultimately hindered by the reliability and completeness of the known functional lists to which comparisons are being made.

(B) Thermodynamics/Secondary Structure Analysis: Thermodynamic/Secondary Structure Analysis is a category broadly encompassing algorithms that attempt to minimize the potential for off-target effects by computationally assessing the binding strength of each putative siRNA to all possible 3' untranslated regions (3' UTRs). siRNA sequences that minimize favorable thermodynamic or structural interactions with transcripts other than the intended on-target are then chosen for use in the screening library (examples: (Naito and Ui-Tei 2012; Chen, Liu et al. 2013)).

(C) Post-Screening Identification of Off-Target Effects: Many bioinformatics approaches have been developed which attempt to identify and remove off-target siRNAs from among already accrued data (also referred to as common seed analysis (CSA)). Typically, these algorithms function by identifying seed sequences which are statistically enriched in siRNA pools responsible for positive phenotypes (Shao, Tsherniak et al. 2013; Zhong, Kim et al. 2014). Particularly sophisticated iterations of this concept include an additional component whereby particular pathways are then integrated into the analysis pipeline to find pathways or processes overrepresented in the positive results (examples: (Chatterjee-Kishore 2006; Sudbery, Enright et al. 2010; Bhinder and Djaballah 2012; Buehler, Khan et al. 2012)).

CHAPTER TWO: Materials and Methods

2.1 Cell Lines

Most human lung cancer cell lines used in this study were established in the John D. Minna and Adi F. Gazdar laboratories. All other cell lines were kind gifts from collaborators as denoted below. Cell lines were grown in RPMI 1640 (Life Technologies Inc) with 5% fetal calf serum (FBS) (or in a few cases 10% FBS for <10% of cell lines). Normal human bronchial epithelial cells (HBECs) were immortalized with ectopic overexpression of CDK4 and hTERT and were maintained in KSFM with supplied supplements (Invitrogen). Some HBEC strains were adapted to growth in serum containing media which was accomplished through the incremental replacement of KSFM with R5 over approximately five to ten passages. Incubation was done in humidified incubators (NuAire) at 37°C at 5% CO_2. Before use and periodically during culturing, every cell line reported in this study was fingerprinted using the PowerPlex 1.2 kit (Promega) and confirmed to be the same as the DNA fingerprint library maintained either by ATCC or the Minna and Gazdar labs. At the same times as fingerprinting, cells were additionally mycoplasma tested (Bulldog Bio). Experiments were performed with cells at approximately 80% confluency excepting when testing conditions required alternate cell numbers.

Cell lines beginning with “H” were established National Cancer Institute and cell lines beginning with “HCC” were established at The University of Texas Southwestern

Medical Center Hamon Center for Therapeutic Oncology Research while colon and pancreatic cell lines were kind gifts from Shay and Brekken Lab on campus. Some of the breast cancer cell lines used were gifts from the Fuqua lab at MDACC or the Martinez lab at UTSW. Cell lines used in the screen can be found in Appendix A.

2.2 Basic Tissue Culturing Techniques

Standard tissue culturing practices were followed for the basic maintenance of the cell lines used in this study. Some examples of such practices include the following.

When passaging cell lines, the media was aspirated, the cells were washed with phosphate buffered saline solution (PBS) and subsequently incubated in a 0.25% Trypsin-EDTA solution (Life Technologies) until cells detached from the plastic. Trypsin was neutralized either through the use of serum containing media or, in the case of HBECs grown in KSFM, Trypsin neutralizing solution (Invitrogen).

Frozen stocks of the cell lines were produced and stored in liquid nitrogen cold storage by first releasing the cells from the plastic as above. Cells were then centrifuged (1000 rpm, 5min), washed in PBS, re-centrifuged, and finally suspended in a mixture of 90% serum and 10% DMSO (Invitrogen).

Stereological pipettes and tissue culture plastic-ware (Corning) were used in addition to biosafety level II containment cabinets. Further basic tissue culture methods can be referenced here (Phelan 2007).

2.3 Transient transfections

All transfections performed were at 20nM siRNA unless otherwise specified. Oligo libraries were purchased from Qiagen as pools of four siRNAs. Half-way through screening, this library was re-ordered from Dharmacon with experimental results demonstrating equivalent results from either synthesis of the library (not shown). Additional oligos were purchased from Sigma for follow-up studies.

Cell lines were optimized for transfection conditions in both 6 well and 96 well formats by varying lipid volume and cell number, and measuring the proliferative difference between scrambled control (Dharmacon) and toxic control oligo PLK1 (Sigma). Because so many cell lines were to be screened in the 96-well format, a detailed, high-throughput methodology was developed to allow the accurate determination of transfection conditions for each cell line. Briefly, eight possible conditions were tested (2000, 4000, 6000, or 8000 cells per well and 0.2 or 0.4 uL of RNAiMAX transfection reagent (Invitrogen) per well) in combination with cells only, lipid only, scrambled control siRNA, and toxic control siRNA controls. Both transfection efficiency (inferred from the killing efficiency of the toxic control) and transfection toxicity (assessed by the comparison of the lipid control and the scrambled control to the cells only control) were considered in choosing the most appropriate transfection condition for each cell line tested.

For 6 well experiments, 2-5 uL of RNAiMAX was added to 0.5mL of serum free RPMI-1640, SFM, and incubated at room temperature for 5 minutes.

A small amount of concentrated oligo was then added to the lipid mixture and allowed to incubate for 20 minutes at room temperature to allow the lipid-oligo complexes to form. Typically, $1.75x10^5$ cells were seeded except when noted otherwise for a final volume of 3mL and a final serum concentration of 5% in 1640 RPMI. When transfecting HBEC cell lines, all parts of the transfection protocol were carried out in KSFM + supplements (for a final volume of 3mL of 100% KSFM).

For 96 well assays an optimized protocol was developed. siRNAs were plated into 96-well plates using UTSouthwestern's High-Throughput Robotics Core Facility at a concentration of 100nM in 20uL of RNAse/DNAse free water (Ambion). Edge wells were not used as testing wells to avoid effects caused by evaporation of the media.

A solution of either 0.2 or 0.4 uL of RNAiMAX transfection reagent (depending on the predetermined transfection conditions (see above)) per 20uL of RPMI 1640 serum free media was made and allowed to incubate at room temperature for 5 minutes to allow for lipid complexes to form. Following incubation, 20uL of this mixture was added to each testing well of the 96-well plate using an eight channel handheld repeating pipette gun (BioRad). After allowing for a 20 minute incubation period at room temperature, varying amounts of cells (2000 to 8000 as determined previously during optimization) were seeded in 50uL of RPMI 1640 + 10%FBS. The final volume of the well was brought to 100uL using serum

free media resulting in a final media condition of 100uL of RMPI + 5% FBS. Cell numbers were determined using a Beckman coulter counter 2000 set at 12um size gating.

2.4 Quantitative RT-PCR

RNA was harvested from cells using the RNeasy Mini Kit (Qiagen). Cells were first washed with cold PBS and then washed directly from the plate using 350ul of lysis solution and collected. A QIAcube automated sample prep system (Qiagen) was then used to reproducibly and accurately carry out the RNeasy Kit protocol. Following protocol completion, RNA quality and concentration was assessed using a NanoDrop 2000 and stored at -80^{o}C.

cDNA was produced using the iScriptcDNA synthesis kit (BioRad)following the kit protocol excepting that 20ul reactions were used and only one fourth the recommended enzyme concentration was used. Target gene TaqMan probes (Applied Biosystems) were used for quantification and internally normalized to GAPDH. RT-PCR reactions were performed in an ABI-7900 Real-time PCR System at 20uL final volume reactions. All reactions were repeated in triplicates and processed in Excel using standard methodology.

2.5 Microarray Analysis

Transcript expression data for all lung and breast cancer cell lines was generated in the Minna Lab using the Illumina (WSG-V3 and V4 BeadChip) array platforms. Tumor cell

RNA was isolated using Qiagen RNeasy kit with genomic RNA filtering and total RNA quality and concentration was confirmed by both NanoDrop 2000 analysis and capillary electrophoresis on the Experion System (Bio-Rad). Total RNA was labeled, amplifiedand re-analyzed for quality prior to hybridization by the UTSW Simmons Comprehensive Cancer Center Genomics Core.

Data was pre-processed using R package mbcb for probe summarization and background correction (Ding, Xie et al. 2008). Following pre-processing data was log-transformed after quartile-normalization. Finally MATRIX (MicroArray Transformation in Microsoft Excel) software version 1.41 was used to import and analyze microarray expression data. Using MATRIX, transcript expression was normalized across samples by the median value, and then normalized expression signals were log_2- transformed and color-coded. Correlations in expression, were determined by Pearson's correlation coefficient. For comparison between sample classes, the ratio of log2- transformed signals from sample classes were generated and two-sample t-tests were performed in MATRIX to filter out non-significant differences in expression ($P = 0.05$). Further analysis was done using the Statistical Analysis of Microarrays (SAM) R package (Tusher, Tibshirani et al. 2001), a permutations based methodology to remove correlations/calculate p-values based on a false-discovery rate.

2.5 Protein Expression

Cell lysates were obtained using M-PER mammalian protein Extraction reagent (Thermo) supplemented with protease and phosphatase inhibitor (Roche) following a standard protocol involving incubation on ice, high-speed centrifugation, and supernatant collection. Lysates for untreated cells were obtained from cells in mass culture collected at approximately 80% confluence. For treated cells, transient transfections using siRNAs were carried out in 6-well dishes seeded at one-hundred and seventy-five thousand cells per well. Multiple wells were then pooled for further analysis.

Protein lysates were quantified using a standard Bradford assay (Promega) protocol and boiled in water for 10 minutes, not a heat block, for even denaturation. Samples were then loaded at a concentration of 50ug per well and then separated using 10% precast SDS/polyacrylamide gels (BioRad) followed by transfer to nitrocellulose membrane (Millipore).Membranes were blocked for one hour at room temperature (RT) in 5% milk solution then incubated at 4°C overnight in primary antibody (see appendix B for overnight directions), followed by secondary antibody (Cell Signaling) for thirty minutes at room temperature. Standard washings in TBSt were performed. Detection of proteins was obtained by enhanced chemiluminescences (Thermo Scientific).

2.7 Colony Formation Assays

Colony formation assays were carried out in 6-well dishes. Cells were plated at low density (100-1000 cells, typically 500 cells) in 3mL of RPMI 1640 + 5% serum. Plates were

shaken according to standard procedures to ensure proper spread of the cells. At reasonable intervals, plates were visually inspected under a microscope to determine the approximate size of the colonies. When colony size reached approximate 50 cells per colony or a different, predetermined endpoint as per the necessary experimental protocol being followed, the assay was considered complete. Following assay completion, plates were stained with a 0.5% crystal violet solution, washed, and allowed to dry. Plates were scanned using a typical desktop computer-attached scanner and quantified both using a "by eye" counting method and a software quantification method. Comparison of results between the two methods gave similar results.

For some experiments, it was necessary to treat the cells being used in colony formation assays. For transient transfections with a colony formation read out, cells were first transfected as normal (see section 2.3) in 96-well plates. Then four hours later, it was first confirmed visually that the cells remained detached from the plate surface. Then the appropriate number of cells was removed (as determined by volume and known plating density) and plated as per normal colony formation protocols. (Note: optimization experiments determined little difference in results between two, four, or six-hour waiting periods though four hours was chosen as a good compromise point between the tested options.)

2.8 MTS Growth Assays

To measure growth rates, relative cell number was quantified using either MTS reagent (Promega). For these experiments, cells were plated and reverse transfected as per section 2.3. Five days following transfection, MTS regent was added to the wells as per the manufacturer's protocol. Plates were allowed 1-4 hours to "turn" the reagent from yellow to medium brown color and then absorbance was measured on a 96-well plate reader and compared to controls to obtain relative cell number information. Likewise, in the instance of drug assays, cells were plated at 500-2000 cells per well and given 24 hours to adhere to the plate. Following the initial plating, drugs were added and then plates were read by MTS assay as per the manufacturer's protocol. Direct comparison was made to Cell Titer Glo reagent using the manufacturer's protocol (Promega) and found equivalent results with both reagents.

2.9 Statistical Processing of siRNA Screening Data

Following transfection in 96-well plates (2.3) and MTS assay (2.8), screening data was further processed in Excel using a visual basic for applications (VBA) macro. Data were transferred into a common sheet, each replicate was analyzed for potential outliers using Grubb's Test with any identified outliers being removed (typically, ~20 data points were removed per screen from a possible 550 data points). Because of the unique nature of the dataset, no standardized methodology existed for the analysis of the data following outlier removal. To ensure accurate interpretation of the data, two different methodologies were

both used and compared for the determination of relative viability and for the identification of hits.

The first of these two methods was to process each individual screening replicate as z-scores. Briefly, z-scores are a statistical look at a dataset that takes into account both the average of the dataset as a whole as well as its standard deviation. Pros to this approach included that it could partially remove differential transfection efficiency as a variable from the analysis as well as allow for more accurate comparison between two different cell lines with largely differing numbers of hits. The major con of this approach was that z-scores were not directly relatable to a tangible concept meaning that the calling of "hits" was arbitrary.

The second method utilized was to compare each siRNA test to the plate average and derive a relative percent viability versus the population average (Note: this was also originally done versus the average of the scrambled control, but when it was discovered that the scrambled control could have significant viability effects (presumably due to off-target effects), this was no longer used). This approach had the major advantage of resulting in an easily interpretable result – percent kill versus control. However, the major disadvantages included the cases of cell lines with many siRNAs resulting in small kills being inaccurately classified as hits. Through comparison of both z-scores and relative viabilities, hits were identified much more accurately.

Finally, outlier siRNA screening replicates were identified and removed utilizing a simple clustering methodology (Section 2.11). Using all the siRNA data, clustering analysis

was performed and it was checked to see if the replicate siRNA screens clustered with one another and were distinct from all other cell lines. In cases where replicates did not cluster, an additional siRNA screen was performed to replace the faulty one.

2.10 Correlation Heat Maps

Correlation heat maps were generated using Excel. Pearson correlations were calculated for all possible pairwise combinations of the data (z-scores, to allow for accurate cross-cell line comparisons) and correlation coefficients were color coded based on a scale. Average correlation was calculated both for inter-replicate comparisons and extra-replicate comparisons.

2.11 Clustering Analyses

All clustering analyses were completed using a combination of Excel VBA macros and R Statistical Software. Briefly, data in Excel were transferred to R through use of a freeware program StatConnector. Excel macros were used to invoke R. Within the R environment, distance matrices were calculated as Manhattan distances using the dist() function and clustering was determined using the Ward methodology (Ward, 1963) through the hclust() function. Data were then transferred back to Excel again using StatConnector and clusters were plotted.

In some cases, it was necessary to determine not only the clustering of the data, but also to determine the "fitness" of the resultant clusters. To accomplish this, R package fpc,

function clusterboot() was utilized (Henning, 2007). Conceptually, this bootstrapping methodology calculates a Jaccard similarity index for each cluster by iteratively adding random noise to the dataset and determining how much noise each cluster can tolerate and still be resolved.

2.12 Identification of Cancer Specific Hits

To identify cancer-specific hits, both z-score processing and difference versus plate average processing of the siRNA data were considered (see section 2.9 for details). These two metrics were correlated against each other and an approximate "% kill versus control" was estimated for each z-score. Thus it was determined that a z-score greater than -0.8 appropriately represented a kill of 30% or less versus controls, a z-score between -1.6 and -0.8 approximately represented a kill of 30% to 40% versus controls, and a z-score below -1.6 to the minimum z-score of -5 represented a kill of > 40% to up to 90% kill versus controls. Hits which scored significant kills (z-score < -0.8) in more than three of the cancer cell lines tested (n=60) but not in any of the HBEC cell lines tested (n=7) were considered as cancer specific hits.

2.13 Correlation of Cancer Specific Hits with Known Phenotypes and Biomarkers

For each cancer-specific hit, the pattern of phenotype across the cancer cell line panel was compared to the phenotypic pattern produced by screening of 20 chemotherapeutic or targeted agents across the same cell line panel through correlative analyses.

To identify expression, mutational, or copy number variation (CNV) deletion biomarkers, the results of the cancer specific hits were run through the elastic-net analysis pipeline. Legacy expression, mutation, and deletion datasets were available in the Minna laboratory for the 60 lung cancer cell lines screened in the siRNA screen. These data were converted to z-scores. R package glmnet() (Friedman 2010) was utilized to generate sparse solutions of highly correlated biomarkers from each of these legacy datasets resulting in 400 significant biomarker predictions. These predictions were then manually curated to identify a small number of putative biomarkers with high likelihood of explanative power.

2.14 esiRNA: Production and Statistical Analysis

esiRNA were designed and produced following the published methodology without deviation (Kittler, Surendranath et al. 2007). Three esiRNA libraries were utilized, two mirroring the 120-target NR/CoReg siRNA library and one matching the 102-target second siRNA library. Screening methodology and data processing were done the exact same as with the siRNA libraries (see sections 2.3 and 2.9 respectively). Each library was screened on a panel of 20 lung cancer cell lines and data were compared between the siRNA and esiRNA screens using Pearson correlations. qRT-PCR of esiRNA knockdowns were

performed as in section 2.4. Interferon response was assessed using relative IFIT1 and PKR mRNA levels with PolyIC (Sigma) as a control.

2.15 Cell Cycle Analyses

Cell cycle was assessed using Propidium Iodide (PI, Sigma) to label DNA content. Cells were trypsinized, washed, and repipetted while gently vortexing into a suspension of ice cold absolute ethanol and incubated overnight at -20C. Cells were then centrifuged, resuspended in 500uL of PI staining solution (47% PBS (vol/vol), 47% 0.1% Triton X-100 (vol/vol), 5% 1mg/mL PI (vol/vol), and 1% stock RNase A (vol/vol, Qiagen)), incubated at 37C for 40 min, washed in PBS, centrifuged, and resuspended in 500uL PBS, filtered through a 70um cell strainer and then FACS analyzed (SCAN, BD Biosciences) with untreated cells used for gating purposes. To determine the percentage of cells in each phase of the cell cycle, data was analyzed in FlowJo Software using a Dean Jett Fox model.

2.16 TargetScan Identification of Predicted Seed Sequence Binding Sites

TargetScan, a program designed to predict targets of miRNAs (Lewis, Burge et al. 2005), was used to identify putative binding sites for each siRNA seed sequences. UTR files were downloaded from the TargetScan website. TargetScan version 60 was used for the analysis. Seed sequences for the siRNAs analyzed were entered into the miRNA database files for the program and full siRNA sequences were entered into the miRNA context scores database files. First, all putative binding sites were identified by simple sequence matches of the supplied sequences to the UTR files. Then, context scores were computed using known

rules of miRNA binding and the full siRNA sequences to determine the likelihood (0 (worst) – 99 (best)) of each possible binding site being real. Excel macros were used to map the results of the TargetScan analysis to known transcripts.

2.17 6-Well siRNA Viability Assessments

Viability following siRNA transfection was assessed by a number of methods as appropriate throughout the study.

(A) Colony Formation Readout. Described in section 2.7

(B) Cell Counts. Transfections were performed as described in section 2.3. At the end of the assay, cells were carefully trypsinized and pipetted to ensure single-cell suspensions. 50uL samples were taken from each suspension and cell numbers were assessed using a Beckman Coulter Counter set at a 13um size filter. Cell numbers were then compared between treatment groups and controls using t-tests.

(C) Crystal Violet Staining. Transfections were performed as described in section 2.3. At the end of the assay, the media was aspirated and cells were gently washed with PBS. Cells were then stained for 2-4 hours with a 0.5% crystal violet solution, were rinsed with water (not direct spray, but with a basin), and allowed to dry overnight. Plates were then imaged using a desktop scanner (Hewlett Packard).

2.18 Annexin IV Staining for Apoptosis Assessment

Annexin IV antibody and staining kit was obtained from Promega. Assay was performed as per the manufacturer's protocol. FACS analysis was performed using an ARIA cell sorter (BD Biosciences) and results were analyzed in FlowJo.

2.19 Ligand Screening

Ligand screening was accomplished through use of a NanoDrop instrument in the UTSW high throughput screening (HTS) core facility. Previous work had determined appropriate seeding densities for each cell line to be tested in the assay (data available through UTSW HTS Core). Cells were seeded into 384-well plates in R5 media and allowed to adhere overnight. Following plating, drugs were added such that each of the 110 ligands being screened was tested as a 12-point concentration curve beginning at a top concentration of 10uM and following four-fold dilutions. All tests were performed in triplicate. Four days following addition of drug, viability was assessed by cell titer glo reagent. A complete list of the ligands tested and their commercial sources is available in Appendix B.

CHAPTER THREE: NR/CoReg Expression from FFPE Samples

3.1 Introduction

3.1.1 FFPE

Formalin-Fixed, Paraffin-Embedded (FFPE) archival tissues and their associated diagnostic records represent an invaluable source of information on diseases where the patient outcomes are already known, particularly cancer. FFPE samples are long-lasting and can even represent patient cases that are no longer seen clinically or obtainable due to changes in medical practice and technology (With, Evers et al. 2011). FFPE archiving is the

most common methodology used worldwide for tissue storage. Unfortunately, although the FFPE process preserves the tissue integrity, it also causes extensive damage to nucleic acids stored within the tissue (Farragher, Tanney et al. 2008).

Extensive work has been done to modify nucleic acid extraction methods from FFPE samples such that the resultant materials are viable for standard assay technology (i.e. qRT-PCR, microarray, RNAseq, etc.) with mixed results (Dedhia, Tarale et al. 2007). Certainly, no standard methodology has yet to emerge, and competing technologies advanced from Illumina, NanoString, Life Technologies, Qiagen, and others continue to be tested for efficacy and cost effectiveness (Gnanapragasam 2010).

3.1.2 Survival Signatures in Lung Cancer

One of the current major goals in lung cancer research is the development of prognostic expression signatures capable of predicting good and poor clinical outcomes for both early and/or late stage patients (example: (Larsen, Pavey et al. 2007)). Unlike other cancer fields (such as breast cancer), no standard molecular definition of lung cancer has been achieved, and prognostic signatures have yet to converge on common gene targets or repeating results (Meta-Analysis (Subramanian and Simon 2010). Study-by-study analysis of the hundreds of published survival signatures is beyond the scope of this document, but excellent reviews have been published (Boutros, Lau et al. 2009; Burotto, Thomas et al. 2014). Overall, these signatures have largely failed to achieve wide-spread clinical

application for a plethora of reasons. Chief among these reasons is that no group has yet to demonstrate a clinical benefit to differential treatment based on the results of a molecular profiling signature. Stated simply, patients will receive the same treatment regardless of whether or not they possess certain signatures, and thus clinicians do not see benefit to subjecting patients to the cost and procedures necessary to classify them.

As a further hurdle, prognostic expression signatures identified in preclinical systems must be identifiable and workable in the much more difficult and clinically-relevant medium of FFPE samples before wide-spread application can be achieved. Previous work from the Minna lab has successfully used Affymetrix U133 plus 2.0 arrays to identify an eleven-gene prognostic signature in FFPE, but unfortunately microarray technology is currently not a clinically tractable methodology and further work is now underway to try and recapitulate these results using a more translatable technology (Xie, Xiao et al. 2011).

3.1.3 NRs/CoRegs as Biomarkers in Lung Cancer

As a class, nuclear receptors (NRs) and their Co-Regulators (CoRegs) have been largely unstudied within the context of lung cancer. Scattered reports have examined the roles specific receptors might play within specific contexts (examples: VDR ligands (Kim, Chen et al. 2012) or PPARg ligands (reviewed: (Li, Lee et al. 2006)) as anti-proliferative agents), but their function as a molecular class in lung cancer is not clear. NRs are especially

interesting to study as a class because of the widespread availability of FDA-approved NR ligands (discussed in detail in chapter 4).

Early studies in the Minna laboratory laid the proof-of-concept groundwork for looking as nuclear receptors as a class in lung cancer by analyzing the expression of the 48 nuclear receptors across a panel of 30 matched tumor normal fresh frozen specimen by qRT-PCR, and then leveraging this dataset to develop a prognostic signature which was then validated in two independent datasets (Jeong, Xie et al. 2010). Further analysis of this dataset revealed its ability to differentiate between non-small cell lung cancer specimen, small cell lung cancer specimen, and normal samples. Some nuclear receptors were identified which exhibited histotype-specific expression, and a few nuclear receptor ligands including AR and ERb targeted compounds were demonstrated to have anti-proliferative effects in cells which exhibited high expression of these receptors (Jeong, Xie et al. 2012). Recently, it was further shown that some nuclear receptor ligands can act synergistically with tyrosine kinase inhibitors in lung cancer (Wairagu, Park et al. 2014).

3.2 Aims and Goals

Based on previous work in the Minna laboratory, the following goals were developed.

(A) Define nuclear receptor expression profiles on a larger sample set to gain statistical power using a clinically tractable methodology.

(B) Define nuclear receptor co-regulator expression profiles using this same sample set and methodology.

(C) Recapitulate the prognostic signature identified in fresh frozen samples using the FFPE dataset.

(D) Leverage the dataset to facilitate biological findings.

3.3 Results

3.3.1 qNPA Technology and Selection of Housekeeping Genes

Quantitative Nuclease Protection Assay (qNPA) technology, developed by HTG Molecular Diagnostics (Tucson, AZ), was developed for the methodology of choice for assessment of NR/CoReg expression from FFPE materials (first reported: (Martel, Botros et al. 2002)). qNPA utilizes a simplistic protocol whereby low amounts of starting material (FFPE or otherwise) are directly lysed in 96-well plates (Figure 3.1). Each well contains programming linkers directly arrayed and printed into the bottom of the well. Specific probes for the genes of interest are designed and linked to the programming linkers. RNA released from the lysing step is bound by probes and S1 nuclease is used to digest any unbound RNA. The remaining RNA is released and binds to the programming linkers and is quantified using imaging software and horseradish peroxidase (HRP) linked probes. Literature reports have shown qNPA technology is comparable to microarray technology (Roberts, Sabalos et al. 2007), accurate (Bourzac, Rounseville et al. 2011), clinically applicable (Erickson 2012), high-throughput (Kris, Felder et al. 2007), capable of producing

survival signatures (Rimsza, Leblanc et al. 2008; Katkoori, Shanmugam et al. 2012), and capable of mediating the discovery of biological findings (Altar, Hunt et al. 2008; Pechhold, Stouffer et al. 2009; Rimsza, Wright et al. 2011; Gerson, Maddula et al. 2012).

Before qNPA technology could be utilized for NR/CoReg expression assessment, a scheme needed to be developed to identify proper housekeeping genes relevant to the lung and breast cancer disease context. To accomplish this goal, microarray data from various sources (n > 1600) was compiled and analyzed to identify genes with low, medium, and high expression with minimum variation across the panel (Figure 3.2**)**. From all available probes, the 92 most promising genes were then placed into two qNPA arrays (EC1 and EC2) and 889 FFPE samples were run across these two EC arrays. Ultimately, five genes (CTBP1, OAZ1, RPSA, EEF2, and RPS19) were chosen as optimal.

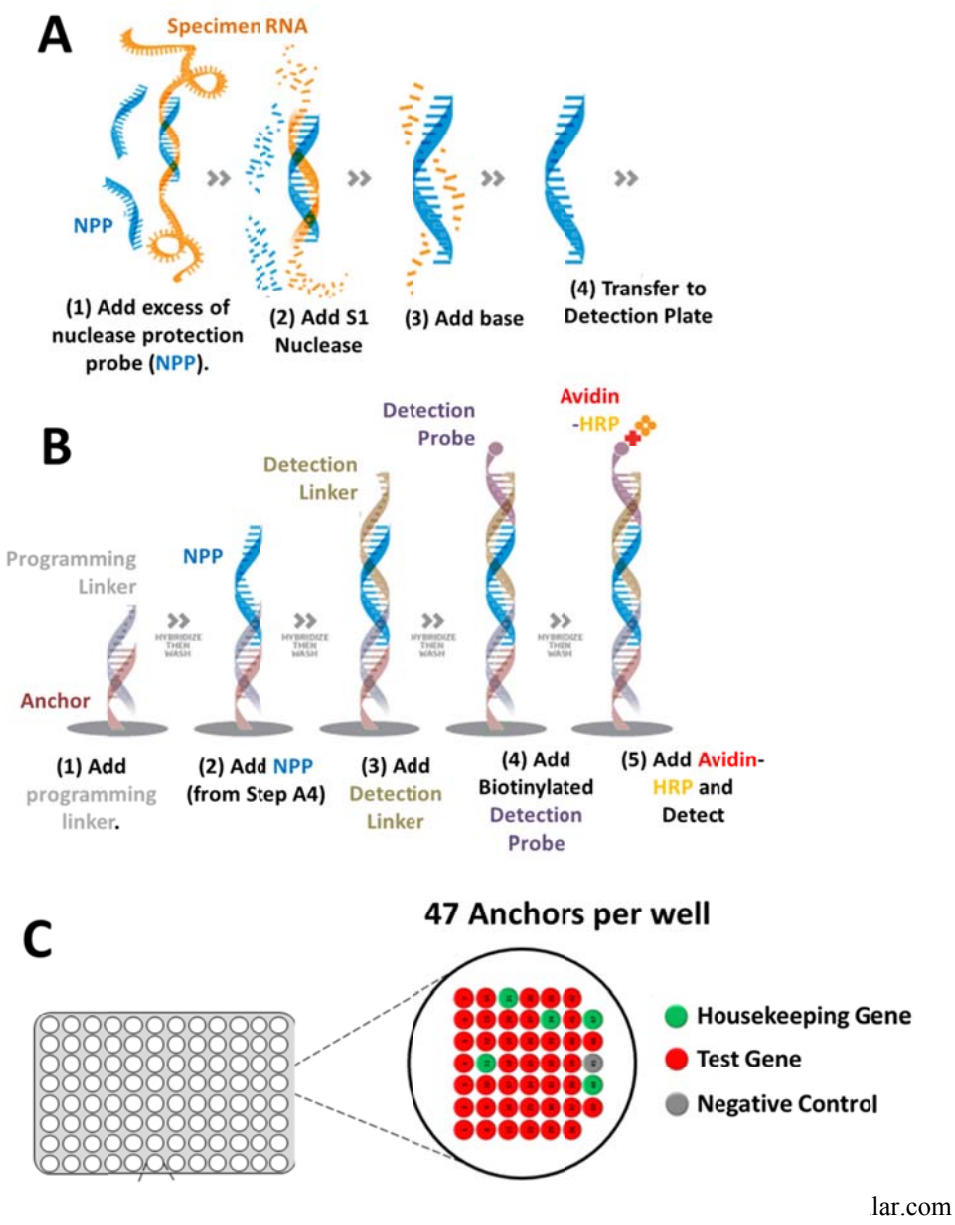


Figure 3.1 Overview of Quantitative Nuclease Protection Assay (qNPA) Process

A) Schematic representation of the preparation of the oligonucleotide library from specimen RNA (released from FFPE by a proprietary lysis buffer).
B) Graphic depiction of the programming linker and Horse Radish Peroxidase (HRP) detection process.

C) Layout of testing plate and housekeeping genes within a single well of a 96-well plate. Multiple samples can be tested simultaneously by using concurrent wells in a 96-well test plate (one sample per well).

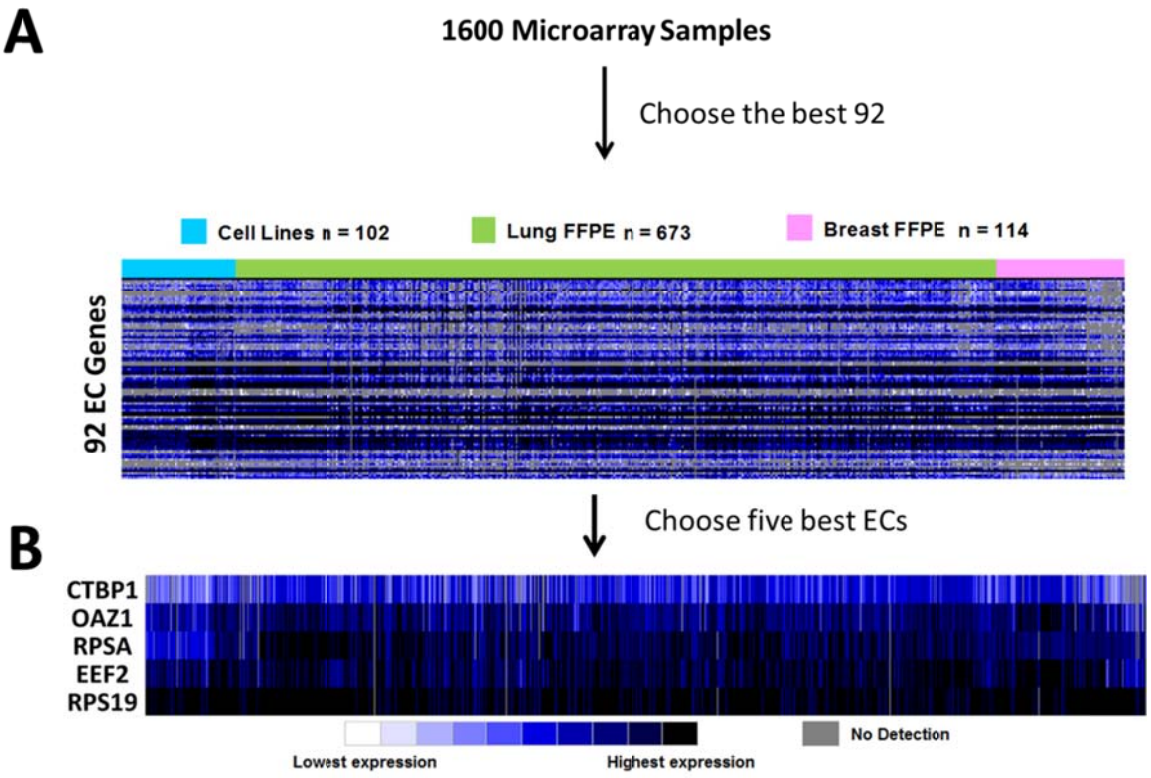


Figure 3.2 Schematic representation of the selection of the five qNPA housekeeping genes.

A) Publically available array datasets were collated and analyzed for genes with low standard deviation and reasonably high average expression across the 1600 sample

panel. The ninety-two best genes were then made into two qNPA libraries and run across both Formalin-Fixed Paraffin Embedded (FFPE) and fresh frozen samples to assess their performance in the qNPA assay.

B) The five best endogenous control (EC) genes were used as quality control (QC) genes in the final Nuclear Receptor (NR) qNPA array.

3.3.2 qNPA Dataset Overview and Quality Analysis

Following EC array development and housekeeping gene selection, a large panel of 850 samples was run in triplicate across the NR/CoReg qNPA system. Average coefficient of variation (CV) between replicates was found to be less than 15%, demonstrating the reproducibility of the technique. Six samples were removed as outliers and cluster analysis of the remaining 844 samples revealed a striking ability of the qNPA dataset to distinguish between the lung and breast cancer samples, the known molecular subtypes of breast cancer (HER2+, ER+, ER-), and between the lung tumor samples and their matched normal counterparts (Figure 3.3).

Basic comparative analyses were performed between the qNPA dataset and other known datasets to assess qNPA's ability to obtain quality data from FFPE specimen. In a previous study, fresh frozen tissue was collected from the exact same 272 lung cancer tumor specimen used in this study. RNA was previously collected from these fresh frozen samples and run on microarrays (data unpublished). Comparison between the fresh frozen microarray data and the qNPA FFPE data revealed approximate correlation coefficients of 0.5 – a strong value for typical cross-platform comparisons. Furthermore, comparison was made between

qNPA data and qPCR data using the exact same RNA extracts. Average correlation coefficients for these direct comparisons from identical starting material were found to typically be greater than 0.7. Overall, it was determined that qNPA results compared favorably with results from other technological platforms.

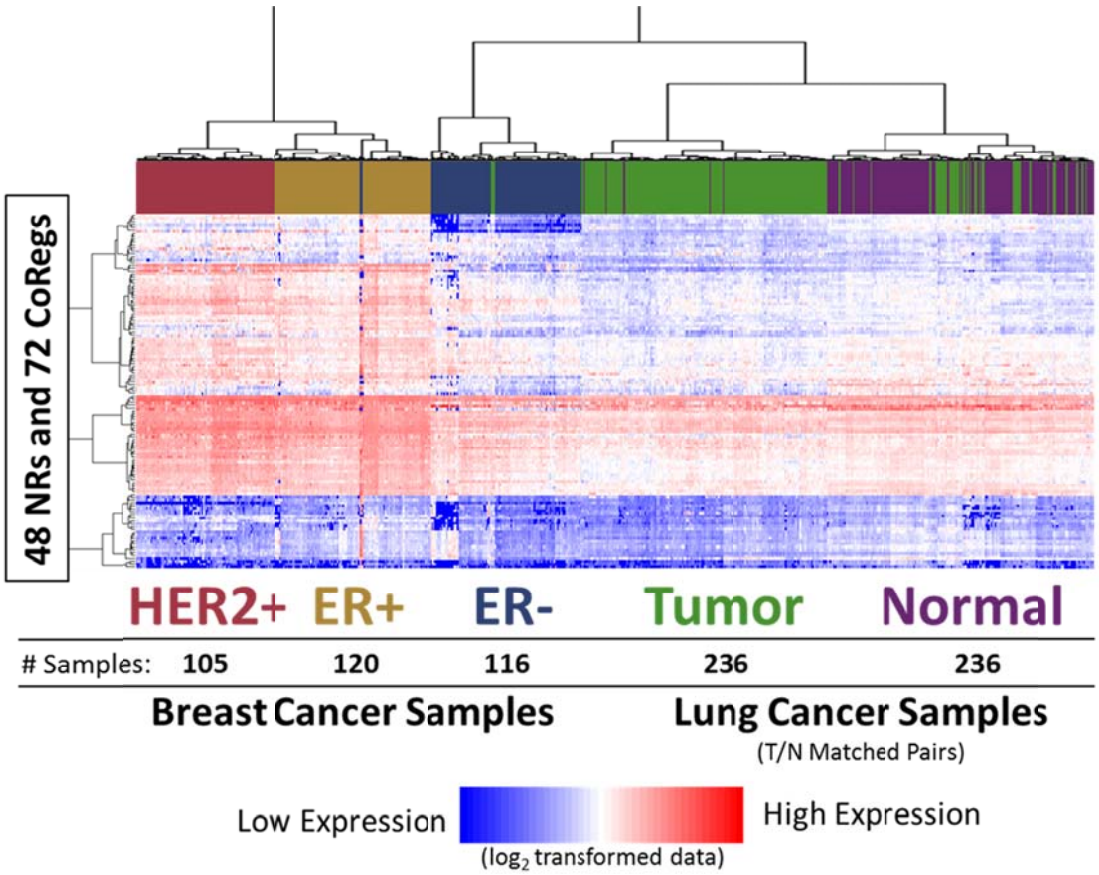


Figure 3.3 Clustering of the FFPE qNPA Dataset.

R statistical software package was utilized to develop the dendrogram. Samples (horizontal axis, top of figure) are colored as per the labels listed along the bottom. Heat map colors represent values on a percentage of rank within the dataset (e.g. the lowest 1% of values in the dataset would be the deepest blue color while genes with the highest 1% of values would be the brightest red color). The approximate range covered between the highest and lowest values is five or six log units – highlighting the strong dynamic range of the qNPA technology.

3.3.3 Analysis of the Breast Cancer Dataset

Unfortunately, because qNPA is not a proven technology, breast cancer (BRC) samples with clinical follow-up data were not able to be secured for analysis in this study. However, basic analysis could be performed on the non-clinically annotated samples that were used here. First, it was necessary to determine how well the HTG ER and PR probes were able to recapitulate the known, clinically-defined levels (Allred Scores) of ER and PR in the analyzed samples. (Allred scores are defined by semi-quantitation of an immunohistochemical (IHC) staining assay, and they are considered the gold standard for clinical definition of hormone receptor positive or negative patients.) In this analysis, it was determined that the qNPA PR probe correlated strongly with the PR Allred Scores and remained linear throughout the necessary range for clinical relevance (Figure 3.4 A). Comparison to ER Allred Score revealed the qNPA ER probe contained sufficient dynamic range to successfully differentiate between ER+ and ER- samples, but lost linearity at the high end of the ER+ range--preventing the differentiation between the highly clinically relevant subsets ER+ (luminal B) and ER+++ (luminal A) (Figure 3.4 B). Future studies will need to adjust the ER range by adding quencher probe to the qNPA analysis wells so all samples can be assessed within the linear range.

Next, sub-clustering analysis was performed to determine whether or not there might be distinguishable subclasses of samples within the molecularly defined ER+, ER-, and

HER2+ subgroups. This question is particularly relevant clinically as patients within ER+, ER- or HER2+ groups have been observed to sometimes have drastically different outcomes. Further subdivisions within these larger groups might allow for increased personalization of therapy or better understanding of risk factors. Within the ER- samples, four subgroups could be identified and defined through expression of three probes: ER, PR, and AR (Figure 3.5 A,C). Approximately half of the samples analyzed fell into the PR+, ER-, AR- subgroup with the remaining samples approximately evenly distributed between ER+, PR-, AR-, and ER-, PR-, AR+, and ER-, PR-, and AR- (Figure 3.5 B). Because the breast cancer samples utilized in this initial study of qNPA technology did not have associated survival outcome data, it could not be determined if these subdivisions were associated with differences in patient survival. Future studies will focus on confirming the existence of these subgroups in independent datasets, and determining if the subgroups can be associated with differences in patient outcome or therapy responsiveness.

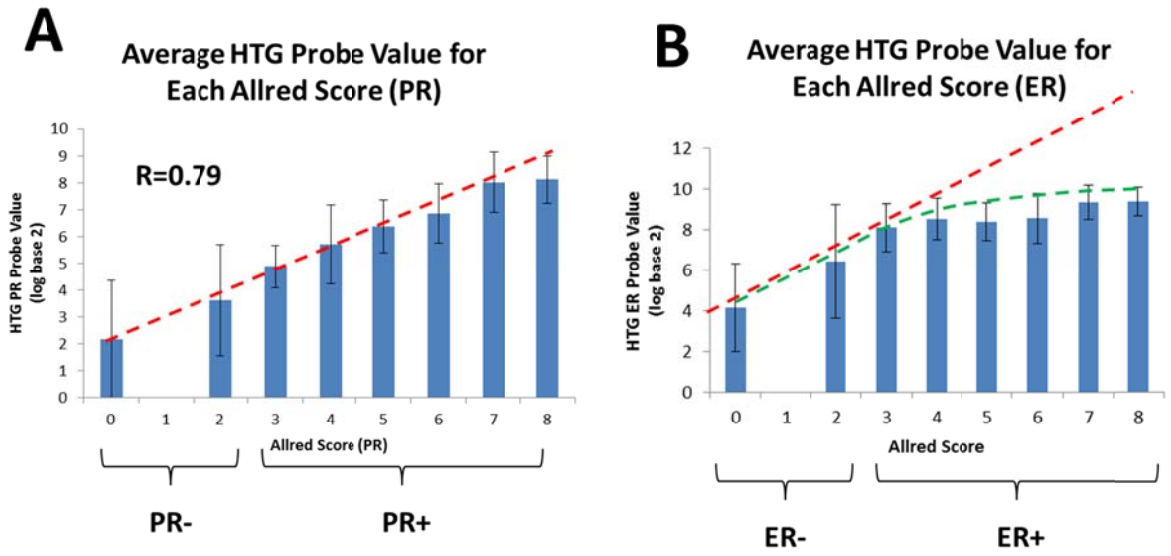


Figure 3.4 Comparison of qNPA and Allred Scores.

A) Samples were categorized based on their clinically assigned Progesterone Receptor (PR) Allred scores (0-8, horizontal axis). Average qNPA PR probe levels were determined for each Allred Score and are displayed as bars on the chart (vertical axis). Correlation between qNPA and Allred Scores was determined by Pearson correlation.

B) Data for the Estrogen Receptor (ER) was analyzed as in (A). The green line was added to approximate the loss of linearity seen with the qNPA probe for samples with high ER Allred Scores.

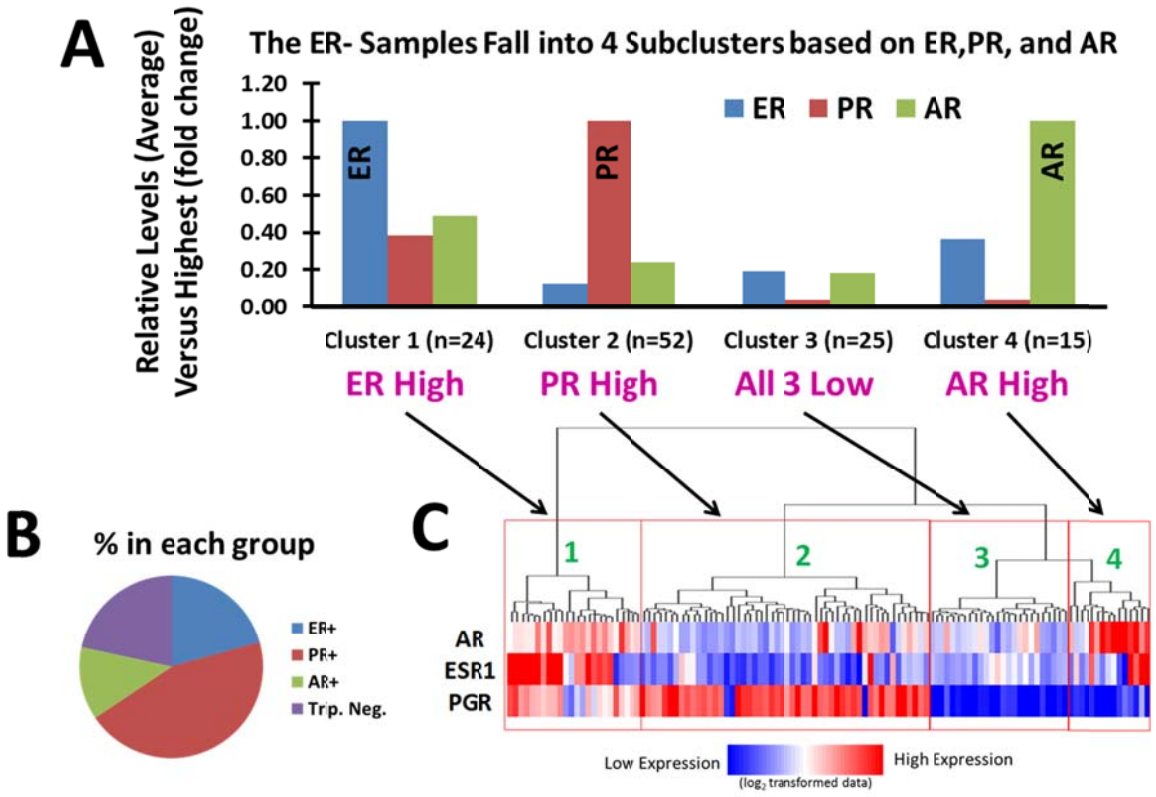


Figure 3.5 Analysis of ER Negative qNPA Data.

A) Relative expression levels of the ER, AR, and PR probes are shown. The cluster with the highest expression of the probes was normalized to 1 to aid in visualization of the results.
B) Visual representation of the percentage of analyzed ER negative samples that fall into each of the four identified subgroups.
C) Clustering analysis of the ER negative samples. All qNPA probes were utilized to generate the clusters but only AR, ER (ESR1), and PGR were seen to be significantly different between the three clusters. This data represents the raw data visualized in (A).

3.3.4 qNPA Re-Finds the NR Prognostic Signature in FFPE Samples

One major goal for the qNPA analysis was to determine whether or not the NR prognostic signature previously reported by the Minna laboratory (Jeong, Xie et al. 2010) could be recapitulated in FFPE samples. Using the same analysis techniques reported previously, the qNPA data were analyzed and the same survival benefit was found whereby patients could be stratified into the same two risk groups albeit with a lower significance than was previously observed (Figure 3.6).

3.3.5 NRs as a Class are Downregulated in Lung Cancer

During data analysis, it was unexpectedly observed that nuclear receptor expression was frequently lost in tumors compared to normal, matched samples (Figure 3.7 A). To corroborate this finding, additional datasets were gathered and analyzed including The Cancer Genome Atlas (TCGA) dataset (Figure 3.7 B) as well as 10 previously published archival datasets available through The Lung Cancer Explorer online analysis tool. Cross-comparison of all these analyses revealed a set of 10 NRs (PPARg, NR4A2, NR4A1, NR4A3, NR5A2, RARb, GR, AR, NR2F1, and RORa) that were consistently downregulated (> 2-fold change, $p < 0.05$) in all the datasets analyzed (Figure 3.7 C). It is important to note that these downregulation events were not the same for each tumor-normal matched pair analyzed, but instead represented the average effect across the dataset. In other words, a

given tumor-matched normal pair could have none of the ten downregulated, some of the ten downregulated, or all ten out of ten downregulated.

Next, survival analyses were undertaken to see if nuclear receptor expression in archival datasets could be a predictor of survival. First, a ten-dataset meta-analysis (n=1405) was performed using the km-plot online analysis tool (Gyorffy, Surowiak et al. 2013). Analysis of all 48 NRs found that patients which retained NR expression had better overall survival than those which lost NR expression ($p = 2.9e-9$, Figure 3.8 A). To confirm this finding, the same methodology was utilized for analysis of the MD Anderson Cancer Center fresh frozen microarray data (n=209). This analysis also found higher expression of the 48 NRs to be associated with better overall survival ($p = 0.0012$, Figure 3.8 B).

These survival analyses were then repeated using the 10 NRs that were found to be consistently downregulated in tumor samples (versus their normal controls). In the ten-dataset meta-analysis (n=1926), loss of the 10 downregulated NRs was found to be strongly associated with worse overall survival ($p = 9.9e-8$, Figure 3.8 C). Approximately one fourth of the samples were identified as still retaining these 10 downregulated NRs while the other approximately three fourths were classified by the algorithm as having lost (on average) these 10 downregulated NRs. The basic distribution of the average expression of these 10 downregulated NRs can be seen in the Beeswarm Graph (Figure 3.8 D).

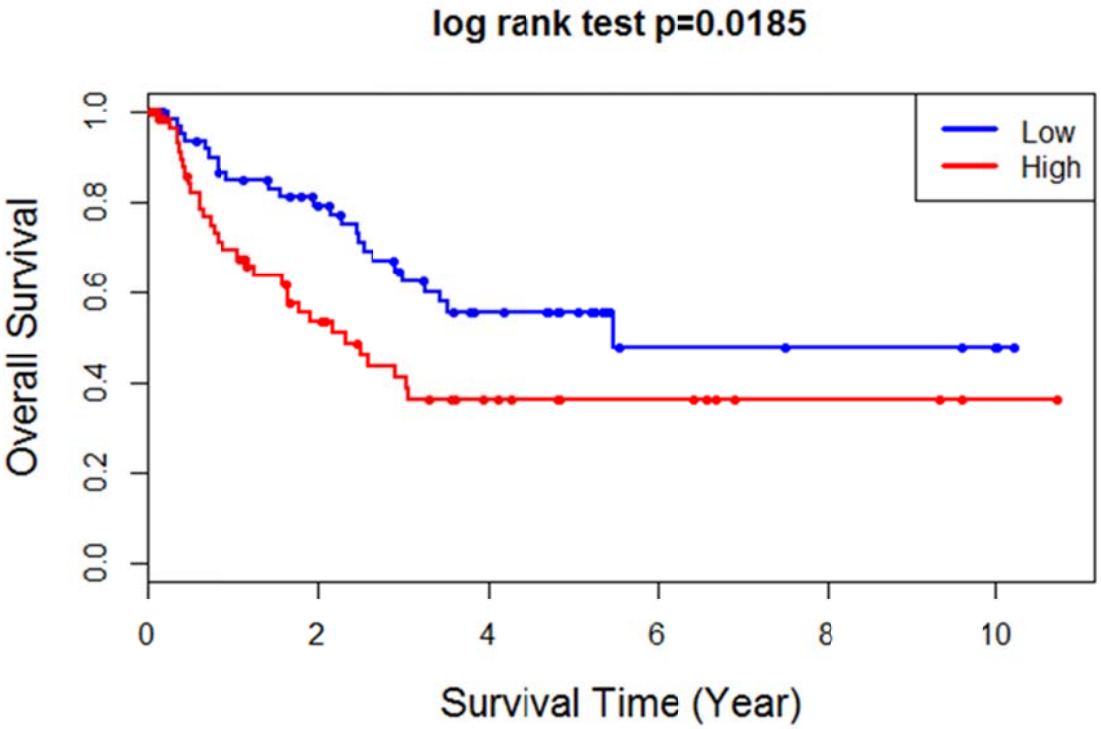


Figure 3.6 Survival Analysis of qNPA Samples

Previously, a survival signature was generated and published which utilized expression of the forty-eight nuclear receptors to predict the overall survival of patients (Jeong 2009). Utilizing the same methodology as that publication, it was determined that the qNPA samples could also significantly ($p = 0.0185$) be subdivided into low and high risk groups based on their NR expression patterns.

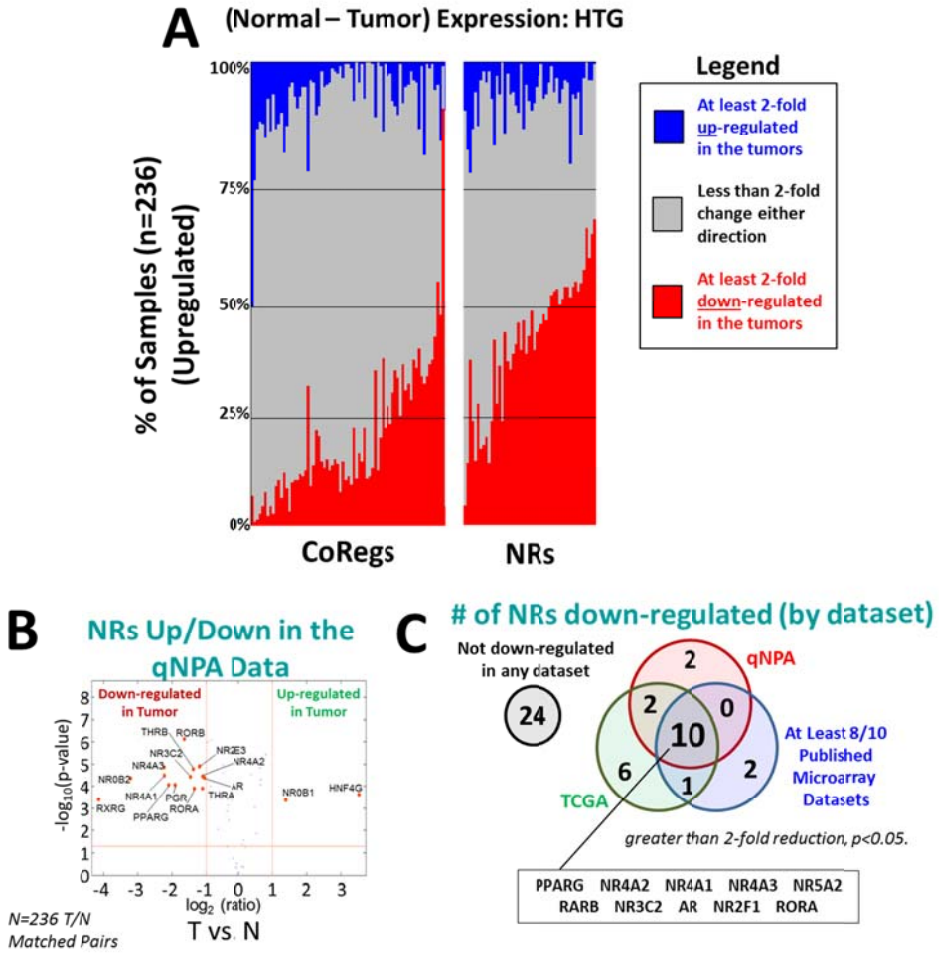


Figure 3.7 A Subset of Nuclear Receptors (NRs) Are Down Regulated in Lung Cancers.

A) Heat map of lung tumor/normal matched pairs data (qNPA, NRs/CoRegs, n=236 T/N pairs). Vertical axis represents the percentage of T/N pairs that have >2-fold upregulation of each NR/CoReg (red). Each column has independently been sorted top to bottom to make data visualization easier. Along the horizontal axis, NRs and CoRegs were each separately sorted left to right from least to most % of samples showing upregulation of the given NR/CoReg.

B) Volcano Plot visualization of the data in panel A. Vertical axis represents –log of the p-value and horizontal axis represents the average fold change between tumor and normal samples. Receptors in the upper left-hand quadrant of the figure (> 2-fold change, $p < 0.05$) are highlighted.

C) Three different sources of data were analyzed to identify a set of ten nuclear receptors repeatedly downregulated in tumor versus normal samples. Specifically, the overlap between the qNPA, The Cancer Genome Atlas (TCGA), and previously published microarray datasets (n=10) is shown.

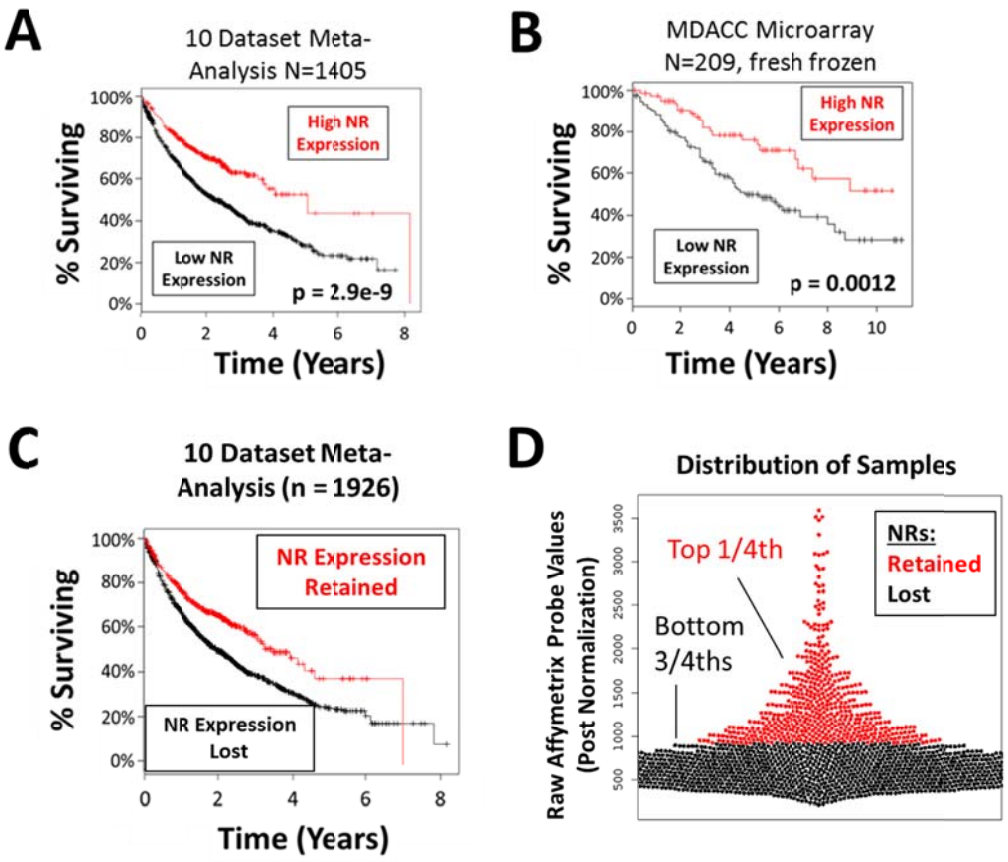


Figure 3.8 Survival Analysis of NRs Expression in Lung Cancer

A-C) Online analysis tool KM-Ploter (kmplot.com) was used to determine survival benefit to loss or retention of NR expression. In panels A and C, analysis was performed on all available lung cancer samples for either (A) all 48 NRs or (C) the 10 NRs identified as repeatedly downregulated in multiple datasets. Probes were weighted equally in the analysis. Panel (B) is the same analysis as in Panel A but for the MDACC 209 fresh frozen microarray samples.

D) Beeswarm plot for the data in panel C. Each of the 1926 samples is represented as a color-coded dot (red or black depending on in which analysis group it was included).

The majority of samples have relatively uniform, low expression of the 10 frequently lost NRs (black) while some samples (approximately 1/4th) have retained NR expression. Higher values on the vertical axis represent proportionally larger average expression of the 10 NRs being analyzed.

3.4 Discussion

In summary, we successfully utilized the qNPA platform to analyze FFPE samples, and utilized the data to successfully make biologically-relevant findings. A panel of 850 FFPE samples was run across the qNPA platform and strong correlation was observed between the qNPA dataset and similar datasets gathered with other technological platforms (microarray, qPCR). Data quality within the qNPA system was assessed and determined to be high with average CVs being around 15% between replicates. Given interest in development and implementation of expression signatures for clinical applications (particularly from FFPE samples), the data presented here strongly support the use of qNPA technologies for simultaneous clinical assessment of expression levels of genetic signatures.

Analysis of the breast cancer dataset revealed successful identification of the generally accepted, molecularly defined subtypes of breast cancer. Furthermore, analysis identified a new sub-classification of ER- samples into four subcategories which could be distinguished by the varying relative levels of ER, PR, or AR expression. Both PR and AR are well known targets in breast cancer therapy, and further subdivision of the much more

deadly ER- breast cancer subtype could lead to new therapeutic options for subsets of ER- patients.

Expression of the 48 NRs and 72 CoRegs analyzed could successfully differentiate between FFPE samples taken from lung cancers versus those from matched controls. It was determined that many NRs were downregulated in the tumor samples versus normal controls in the qNPA dataset. Additional analyses identified ten NRs which were reproducibly lost in tumors versus normal controls in the TCGA dataset as well as in ten additional archival datasets. Furthermore, functional relevance for the loss of NRs in lung cancers was supported by the observation that patients retaining NR expression overall exhibited higher overall survival versus those that had lost expression. These reports represent the first observations of the categorical downregulation of nuclear receptors in lung cancer.

3.5 Future Directions

Based on these findings, it is clear that qNPA represents a strong platform for expression signature analysis. Future studies which aim to identify clinically relevant expression signatures for cancers of any type can be run through the endogenous controls array to identify suitable housekeeping genes. Following housekeeping gene identification, the putative expression signature to be assessed can then be formatted for analysis by the qNPA system.

Further qNPA studies with nuclear receptor expression in breast cancer samples should first focus on ER probe attenuation to ensure that analysis will be within the ER probe linear range for all samples analyzed. Following this improvement to the assay, breast cancer samples with clinical follow-up information available should be run through the qNPA platform to allow for assessment of NRs as prognostic biomarkers in both the current molecularly defined subsets of breast cancer as well as in any sub-clusters that can be identified within each molecularly defined breast cancer subset.

Evidence for a functional role for NRs in lung cancer pathogenesis is strong. Future studies should focus on delineating this functional role through overexpression/re-expression studies to determine what tumor suppressive effect the lost NRs may have been exerting on the tumors. Further analysis into the mechanisms by which tumors downregulate NRs should focus on epigenetic avenues as well as possible therapeutic/small-molecule interventions that might lead to NR re-expression. Particular attention should be given to the ten NRs found reproducibly downregulated across multiple datasets as the strong prevalence of this finding suggests a selection pressure driving this phenomenon.

CHAPTER FOUR: NR Ligand Screening in Lung and Breast Cancer

4.1 Introduction

4.1.1 NR Ligands as Therapeutic Agents

Due to their unique mechanism of action, nuclear receptors (NRs) represent attractive drug targets (see Section 1.1.3). 13% of all FDA-approved drugs target nuclear receptors in a diverse disease set including metabolic, neoplastic, and cardiac diseases in addition to hormonal imbalance syndromes (Overington, Al-Lazikani et al. 2006). Nuclear receptors are master regulators of transcriptional programs within the cell and regulate metabolic processes, cell survival decisions, growth or senescence, and cellular signaling among other functions (Huang, Chandra et al. 2010).

In a cancer context, nuclear receptors are some of the most highly studied molecular targets (Kittler, Zhou et al. 2013; Safe, Jin et al. 2014). A bulk of the work has been focused on breast (reviewed: (Conzen 2008)) and prostate cancers (reviewed: (Savoy and Ghosh 2013)) where nuclear receptors have been comprehensively demonstrated to be completely responsible for the driving and regulation of cancerous signaling within these disease contexts. Other cancers such as ovarian (reviewed: (Zhao, Zhang et al. 2013)), endometrial (reviewed: (Grundker, Gunthert et al. 2008)), or hematological malignancies (example: (Harousseau, Attal et al. 2006)), have also been shown to be driven or largely influenced by nuclear receptor signaling.

4.1.2 NR Ligands in Lung Cancer

In lung cancer, varied success has been observed with identifying biological roles for nuclear receptors and with using nuclear receptor drugs as treatment options in this disease context (Nemenoff and Winn 2005) (Beattie, Hansen et al. 1985). Basic findings will be outlined below on a receptor-by-receptor basis.

(A) **Androgen Receptor (AR)** – No major clinical trials have been performed to date with anti-androgen therapies in a lung cancer context. However, many reports exist which demonstrate that lung cancer preclinical models can respond to androgen related compounds (Mikkonen, Pihlajamaa et al. 2010; Jeong, Xie et al. 2012). Also, AR exhibits cancer-specific expression patterns in lung (Kaiser, Hofmann et al. 1996; Mikkonen, Pihlajamaa et al. 2010), and shows significant

crosstalk with important lung cancer pathways including mTOR and EGFR signaling (Recchia, Musti et al. 2009).

(B) **Estrogen Receptor (ER)** – The selective estrogen receptor modulator (SERM) Tamoxifen has been used in lung cancer clinical trials (Perez, Gandara et al. 2003). In cell line models, aromatase inhibitors have been reported effective for some subsets of patients (Weinberg, Marquez-Garban et al. 2005). Additionally, studies have shown crosstalk between ER and EGFR signaling with combination of EGFR tyrosine kinase inhibitors (TKIs) and fulvestrant (an anti-estrogen that reduces ER protein levels) demonstrating preclinical efficacy (Klinge 2012).

(C) **Glucocorticoid Receptor (GR)**–Glucocorticoids have been studied extensively in the context of lung cancer clinically as a pre-treatment for chemotherapy (Sekine, Nishiwaki et al. 1997). Use of GR-targeted compounds as therapeutic agents has been largely restricted to preclinical models with many reports of preclinical efficacy existing (Greenberg, Hu et al. 2002; Liang, Kowalczyk et al. 2014). Changes in GR expression levels mediated by methylation have also been reported (Kay, Schlossmacher et al. 2011), and GR treatments have been tied to mediation of epigenetic changes (King, Trotter et al. 2012).

(D) **Peroxisome Proliferator Activated Receptor (PPAR)** – No clinical trials have been performed using PPAR-targeted ligands as direct therapeutic agents in lung cancer to date. However, many studies have demonstrated roles for PPARs in lung cancer pathogenesis in preclinical settings (reviewed: (Keshamouni, Han et

al. 2007)). For example, pioglitazone has been demonstrated to prevent smoking induced lung tumors in mice (Li, Kong et al. 2012), PPAR agonists have been shown to induce apoptosis in subsets of lung cancer cell lines (Tsubouchi, Sano et al. 2000), PPAR expression has been shown to be altered in lung cancer tumors and cell lines (Inoue, Kawahito et al. 2001), and PPARs have been shown to be key regulators of lung cancer cell metabolism (Skrypnyk, Chen et al. 2014).

(E) **Retinoid Receptors (RXR)** – Several clinical trials have been performed with RXR ligands in lung cancer (reviewed: (Dragnev, You et al. 2013), key references: (Khuri, Lotan et al. 2000; Khuri, Rigas et al. 2001)). Typically, RXR-targeted agents have been found relatively ineffective as single agents, but have been found to be quite effective in enhancing the efficacy of standard chemotherapy regiments in subsets of patients. Many current studies of RXR ligands focus on 1) understanding the mechanisms by which RXR ligands sensitize some NSCLC patients to chemo, and 2) identifying biomarkers which can predict which subset of patients would respond to an RXR/chemo combination therapy.

(F) **Vitamin D Receptor (VDR)** – Some clinical work has been done trying to treat patients with VDR ligands (particularly calcitriol) (example: (Ramnath, Daignault-Newton et al. 2013)). These studies have occasionally observed partial responses, but are typically dose-limited by hypercalcemia (though next-generation VDR ligands (such as seocalcitriol) have been reported to overcome

this side effect to some degree). Much more work has focused on VDR as a prognostic biomarker in lung cancer. A meta-analysis of twenty-five different studies of VDR as a prognostic marker reported that 1) patients with vitamin D deficiency have a worse prognosis, and 2) supplementation of these patients with vitamin D did not demonstrate a benefit (Buttigliero, Monagheddu et al. 2011).

4.2 Aims and Goals

Based on previous work in the Minna laboratory, the following goals were developed.

(A) Assemble a sizeable library of ligand compounds targeting the 48 nuclear receptors with particular focus on both 1) representative coverage of the non-orphan NRs and 2) ligands which have been used in previous cancer studies.

(B) Perform viability screening of these compounds across a comprehensive panel of lung and breast cancers.

(C) Identify compounds with selective killing effects on specific subsets of the cell lines and identify clinically-actionable biomarkers to predict response.

(D) For promising ligands, use preclinical models to obtain the necessary supporting evidence to result in a clinical trial.

4.3 Results

4.3.1 Compound Screening: Overview

A library of 110 ligands was assembled (Sigma Aldrich) with representative coverage of both agonist and antagonist activity (as was achievable) across the 48 nuclear receptors (summary: Figure 4.1B, complete list: Appendix B). Ligands previously utilized in oncology clinical studies or those with extensive preclinical data were prioritized. Paclitaxel was included as a positive control compound for comparison to standard chemotherapy purposes. Compounds were screened using an ECHO Nanodrop device in 384 well plates, and triplicate 12-point dose response curves were generated for each compound with automated data processing used to determine IC50 values (Figure 4.1A, for complete methodology see Section 2.19). The library was screened against a 130-member comprehensive panel of cell lines (cell line list included in Appendix B).

Preliminary analysis of the dataset revealed a wide range of responses to the compounds (Figure 4.1 C). Some ligands such as clotrimazole, an anti-fungal shown to interact with PXR (Sawyer, Brogden et al. 1975), GSK4716, a selective agonist of estrogen related receptors ERRb and ERRg (Kim, Koh et al. 2009), or dexamethasone 21-methanesulfonate, a covalent binder of GR (Richard-Foy, Sistare et al. 1987) were seen to cause strong kill phenotypes in virtually all of the cell lines tested. Other compounds had no effect on cellular viability in any cell line tested. In addition to the ligands which caused similar effects (either kill or no-kill) in virtually every cell line tested, ligands were also identified that elicited effects in some cell lines but not in others (termed “selective ligands”). While most of the selective ligands in the screen did not correlate with each other (i.e. they

did not have their selective effect on the same cell line subset), a few cases were identified where multiple ligands had a selective effects on a very similar subset of cell lines.

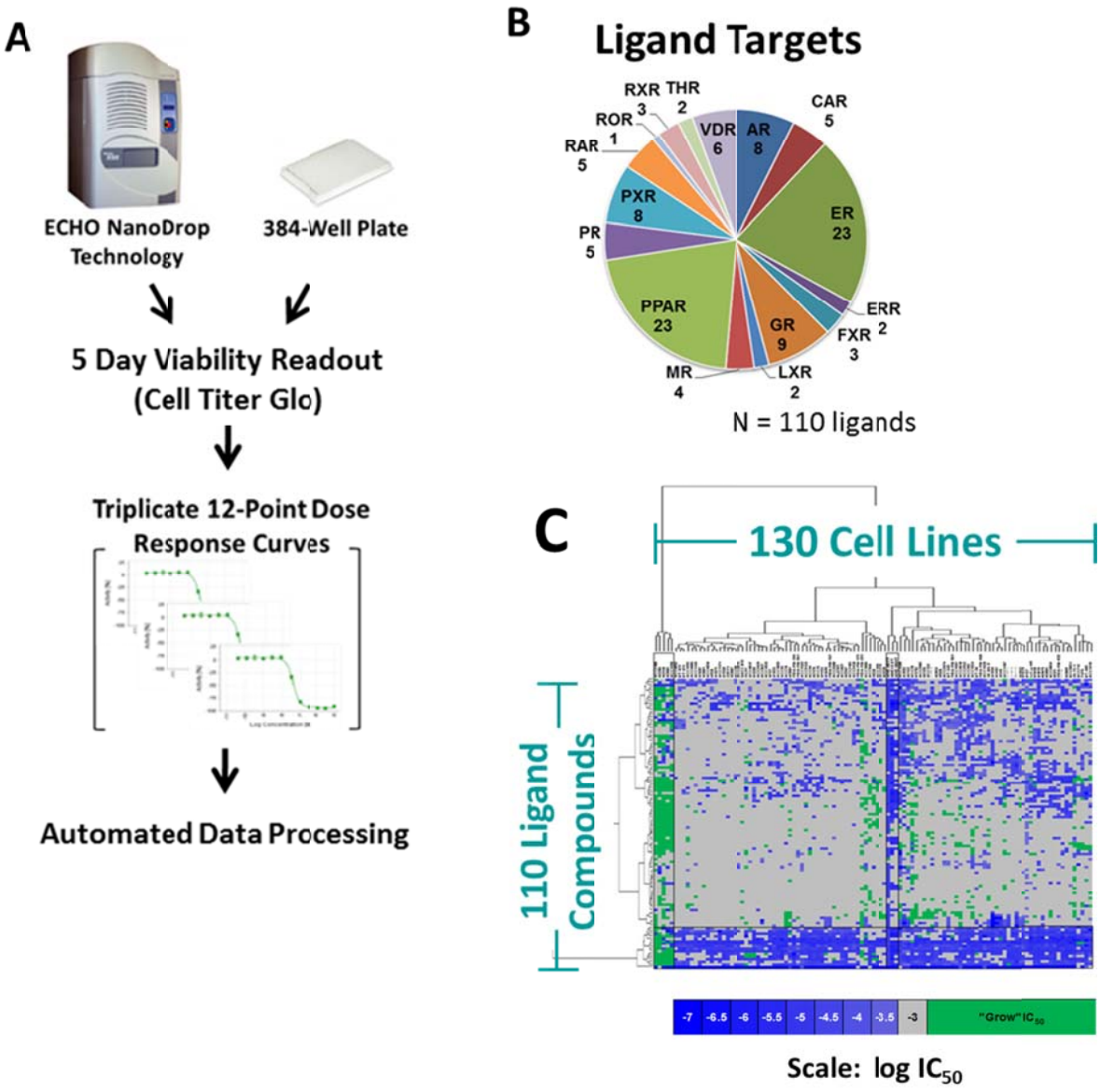


Figure 4.1 Overview of nuclear receptor ligand screening.

A) Schematic representation of the methodology used (for details see section 2.19).
B) Ligands used in the screen shown by the most commonly accepted or referenced target for each ligand. It is important to note that many ligands are known to act upon multiple receptors within the tested concentration range utilized by this screen.
C) IC50 values were determined for 110 ligands across 130 cell lines. Results are color coded as per the scale.

4.3.2 Analysis of Select Compound Classes

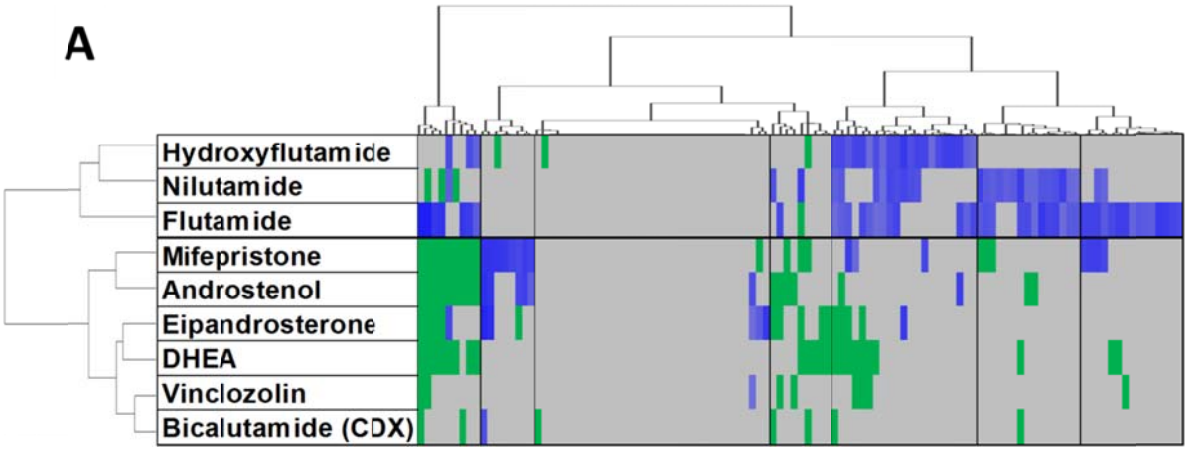
To take advantage of having multiple compounds targeting the same receptors with known mechanisms of action (a unique feature of the ligand library), subsets of the nuclear receptor ligands were analyzed independently for activity across the panel. Several examples of these types of analyses will be briefly highlighted below. However, it is important to note that many ligands are known to target multiple receptors depending on the concentration being considered. Because of this, these analyses were used as an initial guideline with follow-up verification of ligand target and mechanism of action planned for particularly promising analysis results.

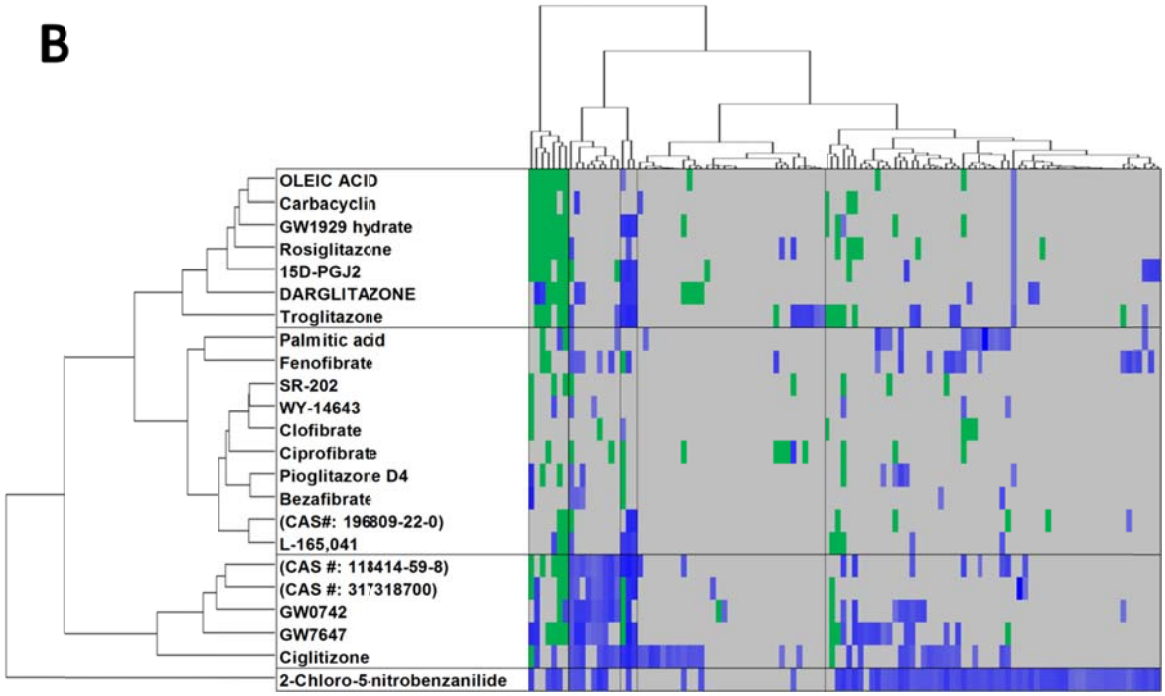
(A) **Androgen Targeted Compounds** – Of the 110 ligands in the library, nine had been previously reported in the literature to interact with the androgen receptor (note: this does not exclude the possibility that they might also interact with other cellular targets). Interestingly, all the compounds tested that were known to be anti-androgens (with the exception of bicalutamide which may have had little effect due to its low solubility) caused viability loss in somewhat non-overlapping subsets of the cancer lines (typically at higher concentrations). Nuances in differences in the mechanisms of action of various anti-androgenic compounds have been reported (Gillatt 2006) and, assuming AR-related activity is confirmed for these kill phenotypes, future studies might focus on elucidating these differences as explanation for these results.

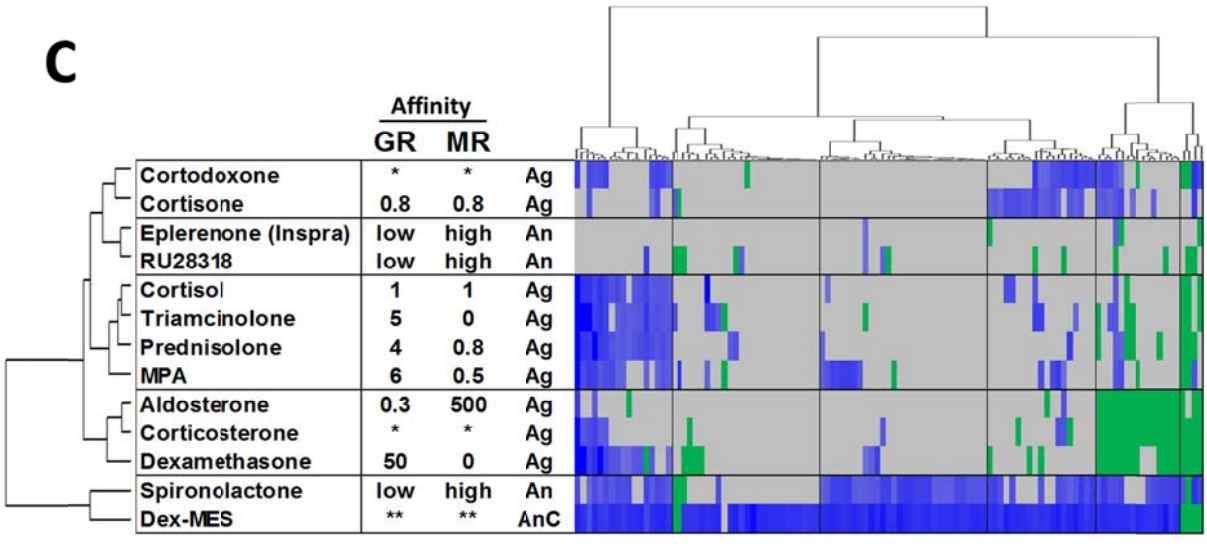
(B) **PPAR Targeted Compounds** – Twenty-three of the 110 ligands in the library had been previously reported in the literature to interact with PPAR receptors (note: this does not exclude the possibility that these ligands might also interact with other cellular targets). Interestingly, similar phenotypic outcomes were noted for specific subsets of the cell lines in response to multiple different PPAR ligands (Figure 4.2B). Unexpectedly, these ligands did not seem to target a specific PPAR, but showed more diversity. However, this finding may be due to the incomplete nature of the literature on the mechanisms of action of some of these ligands than a true finding and further study should be undertaken to confirm the PPAR subtype specificity of these compounds. Interestingly, the compounds available to target PPAR were largely agonists and strong anti-cancer activity was observed with 2-chloro-5-nitrobenzanilide – the only known antagonist in the panel (Abd-Elrahman, El-Gowelli et al. 2010).

(C) **GR/MR Targeted Compounds** – Thirteen of the 110 ligands in the library had been previously reported in the literature to interact with either the GR or MR receptors (these two were considered together as GR/MR ligands are known to frequently interact strongly with both receptors (Hantzis, Albiston et al. 2002)). Known relative strengths and agonist/antagonist status for GR/MR for these 13 ligands are summarized in Figure 4.2C (note: this does not exclude the possibility that these ligands might also interact with other cellular targets). Several interesting subgroups of cell lines could be identified based on their responses to these 13 ligands. First, a

group of cell lines (approximately 10% to 15% of the panel) was identified that exhibited reduced viability following treatment with any ligand known to be at least as strong of a GR agonist as cortisol (triamcinolone, prednisolone, MPA, and dexamethasone). Second, a non-overlapping group was found (approximately 10% of the panel) which showed increased growth following treatment with aldosterone, corticosterone, or dexamethasone. Follow-up studies of this phenotype unexpectedly found that the apparent increase in viability was not due to an increase in actual cell number (data not shown). Because the assay originally used (Cell Titer Glo) to measure cellular ATP levels as a surrogate for cell number, it is hypothesized that these ligands were acting on this subset of the cell line panel such that their metabolic phenotypes were altered. Finally, spironolactone, an MR antagonist (Delyani 2000), was highly effective in a third, non-overlapping subset (approximately 20%) of the cell lines as an anti-cancer agent. Both Eplerenone and RU28318 (also MR antagonists) did not have this same effect, which may suggest an non-MR dependent mechanism for the spironolactone-mediated killing as both have been found to be more MR selective than spironolactone.


A
Hydroxyflutamide
Nilutamide
Flutamide
Mifepristone
Androstenol
Eipandrosterone
DHEA
Vinclozolin
Bicalutamide (CDX)





Scale: log IC_{50}

Figure 4.2 Visualization of the NR ligand screening dataset

Shown are the standalone analyses of ligands demonstrated to target (A) the AR, (B) the PPAR, and (C) the GR/MR. For each case, it is important to note that many of these ligands are known to interact with other receptors as well. Data shown are as IC50 values (as in Figure 4.1). In panel (B), three compounds that did not have common names/abbreviations were indicated by their CAS#s for ease of reading. For panel (C), additional data was shown indicating the agonist (Ag) or antagonist (An) status of each ligand ("AnC" was noted for Dex-MES as this ligand is a covalent antagonist of GR). Furthermore, the strengths of each ligand (relative to cortisol) for GR or MR are shown as determined in previously published work (Miner, Hong et al. 2005). Values not included in this reference are estimated from other publications as "low"/"high." "*" represents a compound which is similar in GR versus MR selectivity to cortisol, but is a weaker agonist. Dex-MES is indicated as an exception (by "**") due to its unique mechanism. Previous reports suggest Dex-MES interacts with both GR and MR.

4.3.3 GR Agonists as Anti-Cancer Agents

Preliminary data analysis of GR/MR agonists revealed that several GR agonists (hydrocortisone (Newman, Nellermoe et al. 1992), triamcinolone (Jeal and Faulds 1997), prednisolone (Gambertoglio, Amend et al. 1980), medroxyprogesterone 17-acetate (Chilvers 1996), and dexamethasone (Sood, Barton et al. 2007)) caused loss of fitness versus controls in a subset of approximately fifteen percent of the tested cell lines (Figure 4.2 C). From these five compounds, dexamethasone was selected as a representative member for further

study due to its widespread clinical use, particular GR selectivity, and current status as a pretreatment for lung cancer chemotherapies (to mitigate side effects).

To begin, experiments were initiated to confirm the cell death phenotype observed in the ligand compound screen. From simple visualization of the cells under the microscope five days post 1uM Dex treatment, it was unexpectedly discovered that dexamethasone treatment did not cause cell death, but instead induced massive morphological changes in the cells (Figure 4.3A). Specifically, the nuclear to cytoplasmic ratio decreased significantly with cells becoming flattened in appearance. Elongated cells were also frequently observed, and multi-nucleated cells could be seen. Further experiments confirmed not only a morphological change in the cells, but also a large decrease in the total number of cells versus untreated controls in a five-day assay (Figure 4.3 B). It was hypothesized that the observed phenotypes might be explained by a cell cycle arrest, and cell cycle analysis was undertaken to test this hypothesis. As expected, a cell cycle block was observed with many of the dexamethasone treated cells arresting in G1 (Figure 4.3 C). Based on these results, it was concluded that treatment of a subset (up to 10% to 15%) of the cell line panel resulted in reduced replication and morphological changes through a cell cycle arrest mechanism.

Following these observations and experiments, studies were undertaken to determine whether or not a biomarker could be found that was capable of predicting dexamethasone sensitivity. Illumina V4 microarray data collected previously was analyzed using Matrix

Software (for details see section 2.5), and a genetic signature consisting of approximately four hundred genes was collated (Figure 4.4 A). The signature was pulled by comparison of the 17 dexamethasone-sensitive cell lines (10 lung, 7 breast cancer) with 39 dexamethasone-resistant cell lines. Unexpectedly, "fold changes" within this signature were drastic, with many genes showing over 100-fold differences between sensitive and resistant cell lines. Interestingly, it was discovered that only the dexamethasone-sensitive lung cancer cell lines exhibited the sensitivity signature while the dexamethasone-sensitive breast cancer cell lines did not. Cell lines that had not been tested in the ligand screen were checked for presence or absence of the Dex signature. Three lung cancer cell lines that had not previously been screened were predicted as dexamethasone sensitive and then screened through the entire ligand panel screen (EKVX, H1944, and H2023). As expected, all three of these cell lines exhibited exquisite sensitivity to the five GR agonists previously shown to be effective in causing cell cycle arrest in lung cancer cell lines exhibiting the dexamethasone-sensitivity signature.

Once the predictive power of the signature had been proven, efforts were undertaken to reduce the signature down to its core component(s). First, comprehensive exome sequencing data (available in the Minna laboratory as an archival dataset for almost all of the cell lines tested in the ligand screen) was analyzed for molecular correlates showing strong association with the dexamethasone sensitivity phenotype. From this analysis, it was determined that cell lines with mutations in LKB1 were statistically overrepresented among

the dexamethasone sensitive cell line group (Figure 4.4 B). LKB1 (STK11), a serine/threonine kinase, regulates cell polarity and is frequently deleted or inactivated in lung tumors (reviewed: (Sanchez-Cespedes 2011). LKB1 is a key player in cellular metabolism and is required for maintaining energy homeostasis and growth under low energy availability conditions. However, it was quickly determined that LKB1 could not be a sole biomarker for dexamethasone sensitivity as approximately 25% of the dexamethasone resistant cell lines were also LKB1 mutants (Figure 4.4 B). To identify a co-biomarker for dexamethasone sensitivity, expression signatures (from available archival microarray data) were compared between LKB1 mutants that were sensitive to dexamethasone and LKB1 mutants that were insensitive to dexamethasone (using Matrix Software). CPS1 was found to be more than 100-fold upregulated in LKB1 mutant, dexamethasone sensitive cell lines versus LKB1 mutant, dexamethasone insensitive cell lines (Figure 4.4 C). Carbamoyl phosphate synthetase (CPS1) is a metabolic enzyme involved in pyrimidine and arginine biosynthesis and in the urea cycle ((Simmer, Kelly et al. 1990)). Few publications are currently available examining the role(s) that CPS1 might play in cancer. Analysis of cBioPortal data (summary site for the TCGA data results, cbioportal.org) showed a relatively high rate of CPS1 mutation (up to 10% - 15% of samples tested, n = 172 samples) in lung cancers, though no recurrent mutations were seen. It is important to note that neither CPS1 nor LKB1 were found to be able to serve as a sole biomarker for dexamethasone sensitivity. As with the LKB1 mutations, cell lines were found in the panel that had high CPS1 and were

dexamethasone insensitive. Only cell lines that had both markers (high CPS1 and LKB1 mutation) could be confidently predicted to be dexamethasone sensitive.

Finally, a preliminary analysis was undertaken to see if a "CPS1-high" subpopulation could be identified in a patient dataset. Analysis of the TCGA expression dataset for lung cancer showed a subset of patients with abnormally high CPS1 expression versus the rest of the analyzed panel (Figure 4.4D).

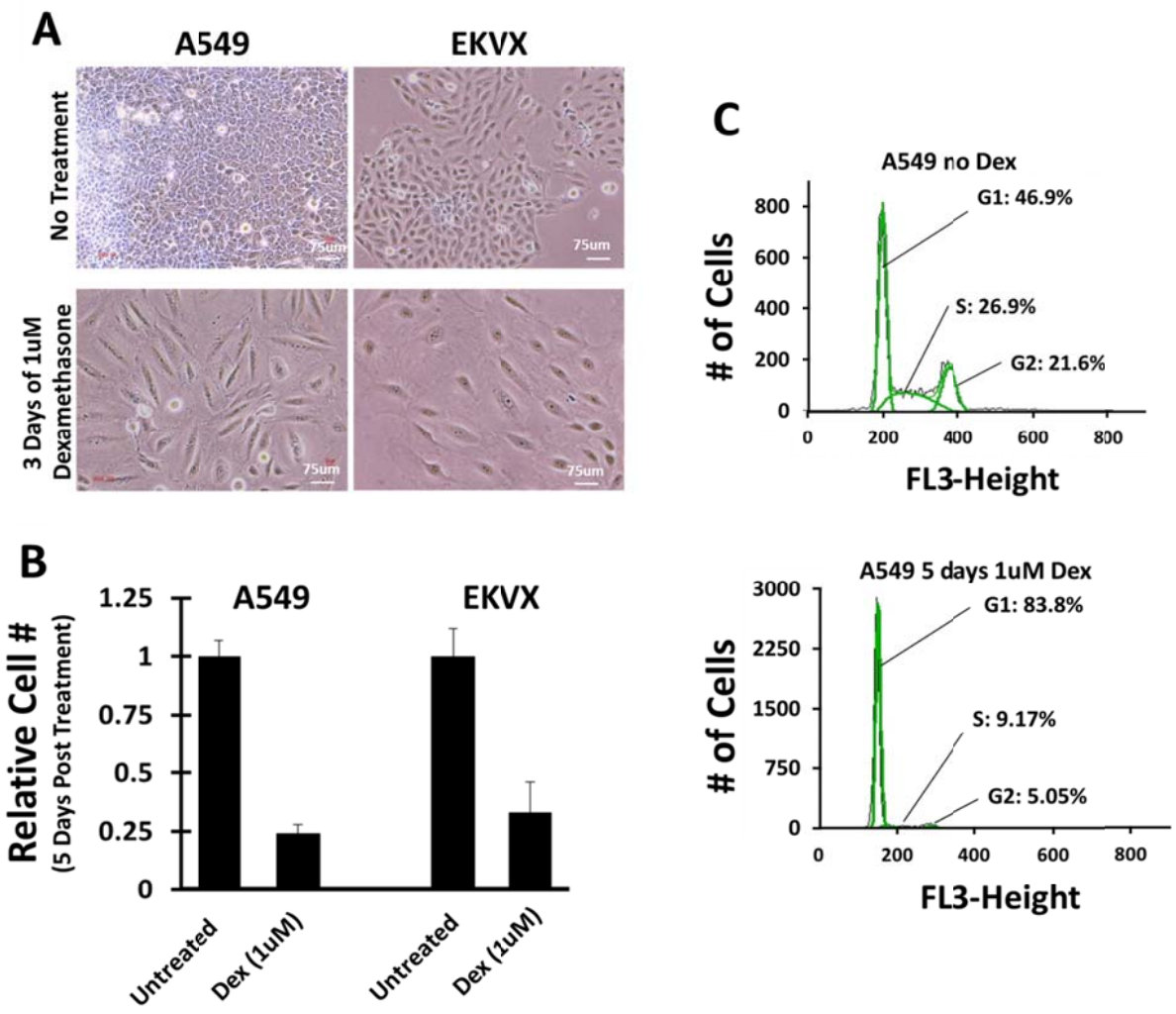


Figure 4.3 Effects of Dexamethasone treatment on NSCLC cell lines
Representative results for Dex treatment are shown with A549 and EXVX. Similar results were seen in the other Dex-sensitive lung cancer cell lines.

A) 10X pictures of two Dexamethasone (Dex) sensitive cell lines (A549, EKVX) five days following 1uM Dex treatment.
B) Five days post 1uM Dex treatment, treated and untreated samples of A549 and EKVX were harvested (trypsin) and counted using a cell counter. Numbers shown are relative to untreated and represent averages of triplicate experiments.
C) Comparative cell cycle profiles for untreated A549 cells (top) and A549 cells treated for five days with 1uM Dex (bottom). 1uM Dex treatment causes a strong shift in the cell population towards the G1 phase of the cell cycle.

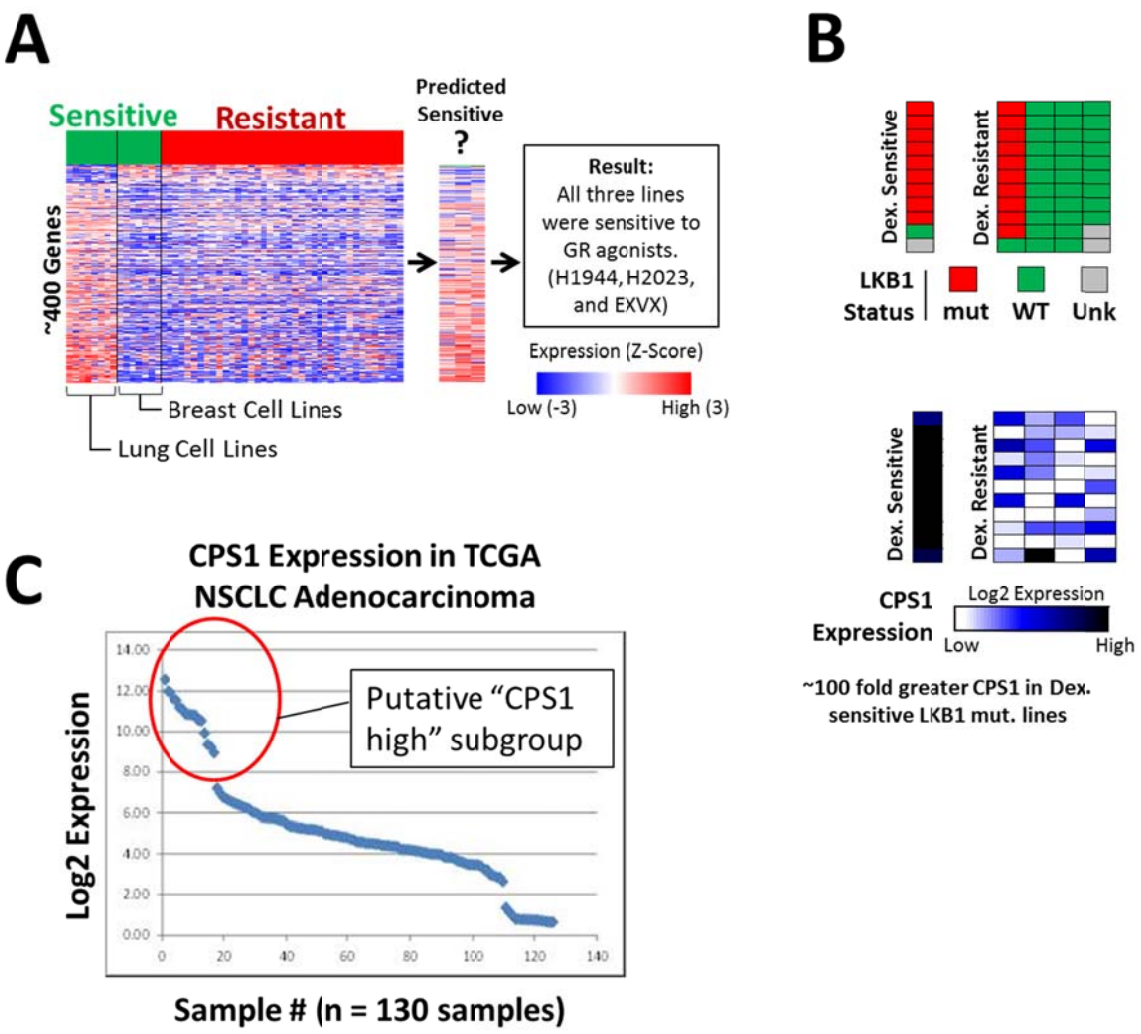


Figure 4.4 Discovery of LKB1 mutation and CPS1 overexpression as a dual biomarker for NSCLC Dex sensitivity.

A) Genetic signature predicting Dex sensitivity. Results are shown as z-scores with the difference between red and blue representing as large as 100-fold differences in expression. The signature was applied to untested cell lines, and was found able to successfully predict three new Dex sensitive cell lines (H1944, H2023, and EKVX).

B) LKB1 mutational status (top) and CPS1 expression levels (bottom) for the NSCLC cell line panel. Cell lines are shown in the same order in both panels.

C) Analysis of CPS1 expression across the TCGA NSCLC-Adenocarcinoma panel (n=130). A subset of samples (circled in red, approximately 15%) show particularly high expression of CPS1.

4.4 Discussion

Nuclear receptor ligands represent a large class of compounds with tremendous translational potential currently not utilized in lung cancer treatment and historically understudied within the context of this disease. To address this issue, a representative panel of 110 nuclear receptor ligands was screened against a comprehensive panel of breast cancer and non-small cell lung cancer cell lines (n=130) fully representative of all currently known molecular definitions of these diseases. Data confidence was exceptionally high with compound screening being carried out under a unique protocol utilizing both 12-point dose response curves and triplicate results being obtained for all pairwise combinations of drugs and cell lines. This dataset provides a strong foundation for future studies to use as a starting point for hypothesis generation regarding the potential uses of NR ligands in lung cancer.

To provide a roadmap for successful analysis of the ligand dataset, receptor class case studies were performed with the Androgen Receptor (AR) ligands, the Peroxisome Proliferator Activated Receptor (PPAR) ligands, and the Mineralocorticoid and Glucocorticoid Receptor (MR, GR respectively) ligands. These analyses revealed that our knowledge of NR ligand functions could be useful in understanding the dataset. For example, a subset of cell lines (approximately 10% to 15%) was identified that responded to all five of the known GR agonists in the ligand panel, which suggests the likely involvement of the GR in mediation of the phenotype.

After analysis of the data, dexamethasone was selected as a hit for further study. Dexamethasone (Dex) has widespread clinical use, exhibits particular GR selectivity, and is currently given to most lung cancer patients as a pretreatment to help mitigate the side effects of chemotherapy. In follow-up studies, it was determined that Dex did not actually cause the active killing of Dex-sensitive cell lines, but instead arrested their growth. Morphologically, Dex treatment had little effect in Dex-insensitive cell lines, but caused distinct changes in Dex-sensitive lines including the enlargement and elongation of cells. Because only a subset of the NSCLC cell lines were found sensitive to Dex (approximately 10% to 15%), efforts were undertaken to identify a biomarker predictive of Dex sensitivity. Initial efforts identified a four-hundred gene signature which was subsequently demonstrated to be capable of accurately predicting whether or not unknown cell lines would be sensitive to GR agonist treatments. Further refinement of this signature ultimately identified an unusual dual biomarker whereby cell lines with both LKB1 mutations and high expression of CPS1 could reliably be predicted to be sensitive to Dex treatment. Both LKB1 (a known tumor suppressor in lung cancer inactivated in up to 20% of patients) and CPS1 are known players in the regulation of cellular metabolism. It will be interesting to see if future studies can further elucidate the link between Dex sensitivity, cancer metabolism, and these two markers.

4.5 Future Directions

The studies presented in chapter four represent only the beginnings of a potential undertaking to identifying novel roles for nuclear receptor targeted therapies in the treatment of lung cancers. As mentioned in the discussion section, the breadth, comprehensiveness, and reproducibility of this ligand dataset are relatively unparalleled in the current lung cancer literature, and many future investigators will likely utilize this resource as a hypothesis generating point for studies on individual compounds or cell line subsets represented in this dataset.

As for more specific future goals, initial efforts should focus on finishing receptor class analyses for the remaining nuclear receptors (such as estrogen receptor, or vitamin D receptor) represented in the ligand library. Such analyses may reveal additional ligand classes with similar mechanisms of action that all give the same phenotypic response pattern across the cell line panel. These could then be quickly studied using a similar pipeline as what was done with dexamethasone.

Other interesting ligands identified in the preliminary analysis of the dataset included Tamoxifen and Seocalcitol. Tamoxifen is well-studied in the context of breast cancer and represents a rapidly translatable compound (reviewed: (Clemons, Danson et al. 2002)). It has been studied as a combination therapy in phase I studies in lung cancer, but detailed analyses to identify clinically actionable biomarkers have been incomplete (Perez, Gandara et al. 2003). In the ligand studies here, it was discovered that many lung lines responded to Tamoxifen treatment at a level comparable to the responses seen from ER+ breast cancer cell

lines (data not shown). Likewise, Seocalcitol, a well-known vitamin-D analogue used in several high profile phase I and II studies in hepatocellular (Dalhoff, Dancey et al. 2003) or pancreatic (Evans, Colston et al. 2002) cancers, showed strong effects in many of the cell lines tested.

To complete the preclinical studies of dexamethasone, it will be necessary to first demonstrate in mouse models that dexamethasone effectively inhibits tumor xenograft growth. Furthermore, it will be essential to examine the potential link between dexamethasone treatment and chemosensitivity. Currently, dexamethasone is given to virtually all lung cancer patients (through a variety of dosing schemes) to help mitigate chemotherapy-induced nausea and vomiting (Peterson, Hursti et al. 1996). In addition, many literature reports exist arguing that dexamethasone can have e a chemoprotective effect (example: (Zhang, Beckermann et al. 2006), a chemoenhansing effect (example: (Wang, Wang et al. 2007)), or no effect in combination with chemotherapy. Most of these studies were done across differing tissue types, with different dosing schedules, and using different model systems. It will be important to establish whether or not dexamethasone treatments enhance or diminish the effectiveness of chemotherapy in lung cancer models. Additionally, further experiments ought to focus on elucidating the potential mechanisms of dexamethasone action in the dexamethasone sensitive cell lines. Potential avenues of exploration include timecourse-based microarray analyses with and without dexamethasone to determine what transcriptional programs dexamethasone might be regulating, chromatin

immunoprecipitation assays before and after dexamethasone treatments to determine how dexamethasone treatment might be effecting glucocorticoid receptor promoter occupancy, and also metabolic assays to explore the potential mechanistic roles that CPS1 might be playing.

CHAPTER FIVE: RNAi Screening of NR/CoRegs in Lung and Breast Cancer Cell Lines

5.1 Introduction

5.1.1RNAi Technology for Cancer Phenotyping

Since its discovery, RNAi technology has been utilized to great effect for a plethora of purposes within in virtually every context of biomedical research. Of particular import, RNAi technology has allowed the systematic interrogation of the functional space surrounding the cancer genome in a timely and efficient fashion (Mohr, Bakal et al. 2010). A variety of approaches have been employed including differing scopes of studies. Many studies utilize RNAi as a secondary technique to validate or investigate findings identified through other methods. More relevant here, a particularly poignant use of RNAi has been in screening efforts. Typically, two variables must be considered in the design of these screens: 1) the number of cell lines to be tested and 2) the number of siRNAs to be screened.

To date, virtually all large scale RNAi efforts performed have focused on interrogating many genes across few cell lines (reviewed: (Sigoillot and King 2011)). In direct contrast, the efforts described here aim to screen a relatively small number of siRNAs

(n=120) across a relatively large number of cell lines (n=130). This unconventional goal presented challenges and also allowed us an uncharacteristically close look at the data generated from a different angle than typical large-scale screens (namely, we were interested in the whole dataset rather than just identifying a single hit), and this afforded us the opportunity to further understand RNAi technology and its potential limitations and applications.

5.2 Aims and Goals

The following list of aims and goals was developed both initially and during the evolution of the project.

Original Goals:

(A) Develop a screening methodology and transfection conditions for medium-scale siRNA libraries being screened across large cell line panels.

(B) Determine the effect (if any) of systematic loss-of-function of the forty-eight nuclear receptors and a panel of seventy-two co-regulators across the cell line panel.

(C) Select one or two particularly promising siRNA pools and further characterize their (originally: "on-target," later: "off-target") effects and potential as cancer-specific vulnerabilities.

Additional Goals:

(D) Characterize the off-target effects in the siRNA screen and add to basic biological understanding of siRNA off-target effects as necessary such that goals "C" and "E" can be accomplished.

(E) Discover the true target(s) of an siRNA off-target effect.

(F) Identify a list of seed sequences that serve as cancer-specific hits.

5.3 Results

5.3.1 Design of the siRNA Screen

Because of the unique needs for the siRNA screen proposed (large number of cell lines, few siRNAs), a new methodology needed to be developed before the dataset could be generated (Overview in Figure 5.1A). To overcome this challenge, a combination of high and low throughput methods were combined. First, siRNAs were plated into screening "daughter" plates using 96-channel automated pipetting systems (Biomek, Beckman Coulter) available in the UTSouthwestern high-throughput core screening facility. This methodology was efficient because each cell line to be screened required the same plate layout template and hundreds of plates could be made in a single day. Following siRNA plating, the daughter plates were sealed and frozen at -20C until needed. Because each cell line to be screened was only to be screened against a total of six 96-well plates, it was not efficient or cost effective to do the remaining portions of the siRNA screen using high-throughput

robotics. Instead, a simple protocol was developed using an eight-channel automatic repeating pipette (BioRad) which allowed for cost-effective plating of the lipid transfection reagent and cells.

Unfortunately, it was also determined that transfection conditions in 96-well settings were particularly sensitive, and it was necessary to thus individually optimize conditions for each cell line that would be screened. To accomplish this goal, a simple methodology was developed which allowed for the simultaneous optimization of six cell lines--again using an eight-channel automatic repeating pipette. Various concentrations of lipid-per-well and numbers of cells-per-well were tested against lipid only, scrambled siRNA, or toxic siRNA conditions, and controls were compared to identify the most effective conditions (Figure 5.2A). It should be further noted that not every cell line tested could be successfully transfected by these methods. Particularly, cell lines which grew without anchoring to the plastic surface (typically small cell lung cancers), cell lines which grew in small clumps or clusters, or cell lines with exceptionally long doubling times were determined to be unable to pass quality control tests and were omitted from the screen. Additionally, it should be noted that cell lines with mycoplasma infections were found to be greatly reduced in their ability to be transfected and to yield reproducible results. As a result, extreme care was taken to ensure that all cell lines were mycoplasma free before their use in these assays.

Although the number of cell lines screened was extensive, not every available cell line could be feasibly screened due to cost and materials restraints. As a result, care was

taken to assemble as representative of a cell line panel as possible (Figure 5.1A). Within the lung cell lines (n=65) care was taken to ensure representation of all relevant lung cancer oncogenotypes (KRAS, LKB1, p53, EGFR, etc.) as well as known cell line pairs (when possible). Cell lines with the most available legacy data (expression data, mutational data, drug sensitivity data, reverse phase protein array (RPPA) data, radiation sensitivity data, copy number data, methylation data, etc.) were prioritized over unknown cell lines. Originally, a panel of fifty cell lines was selected, but some cell lines were found difficult to transfect and were replaced with other cell lines in addition to 15 further cell lines to ensure a comprehensive panel. Unfortunately, the histotypic representation of the panel was not reflective of the clinical situation due to the technical challenge of transfecting small cell and squamous lung cancer cell lines. Within the breast cancer cell line panel, all available breast cancer lines were tested and any that passed optimization protocols were included in the screen (n=24). All relevant molecular disease states of breast cancer (ER+, PR+, HER2+, triple negative) were represented. Additionally, seven parental human bronchial epithelial cell lines (HBEC) were screened along with all possible oncogenic variants of these lines including p53, KRAS, LKB1, c-MYC altered lines and lines with and without growth adaptation to serum conditions. For comparison purposes, seventeen additional cell lines from other cancer types (melanoma, hepatocellular carcinoma, prostate cancer, and pancreatic cancer) were also screened as materials were available.

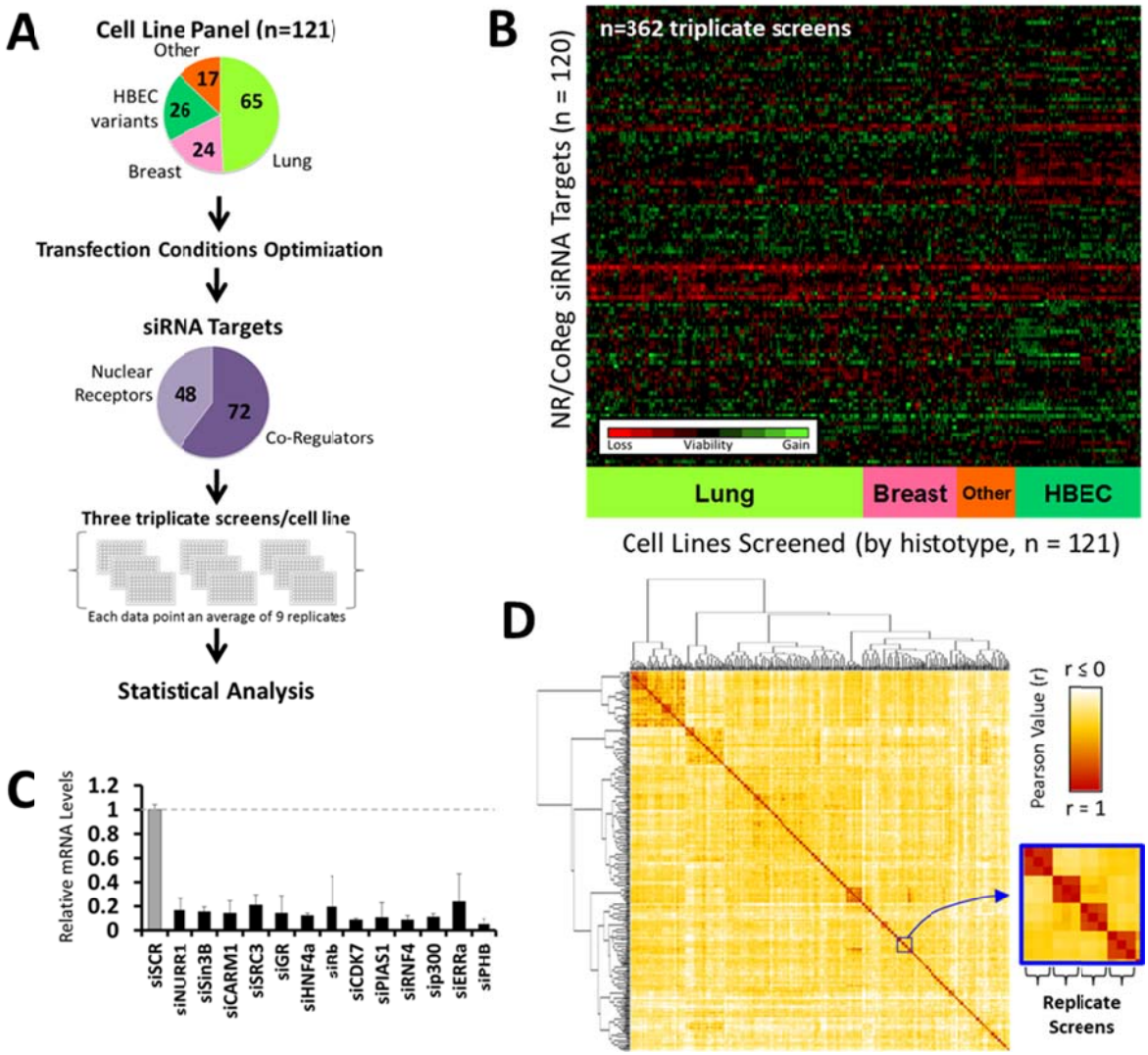


Figure 5.1 Overview of the siRNA screen

A) Schematic of screening methodology.
B) Heat map representation of siRNA screening data (Z-scores).
C) Knockdown efficiencies of 13 siRNA pools. Results represent average knockdown across three cell lines (error bars, SD).
D) Heat map of correlation coefficients (Pearson) between all siRNA screens. The enlargement highlights the high correlation (avg. = 0.863) between replicates.

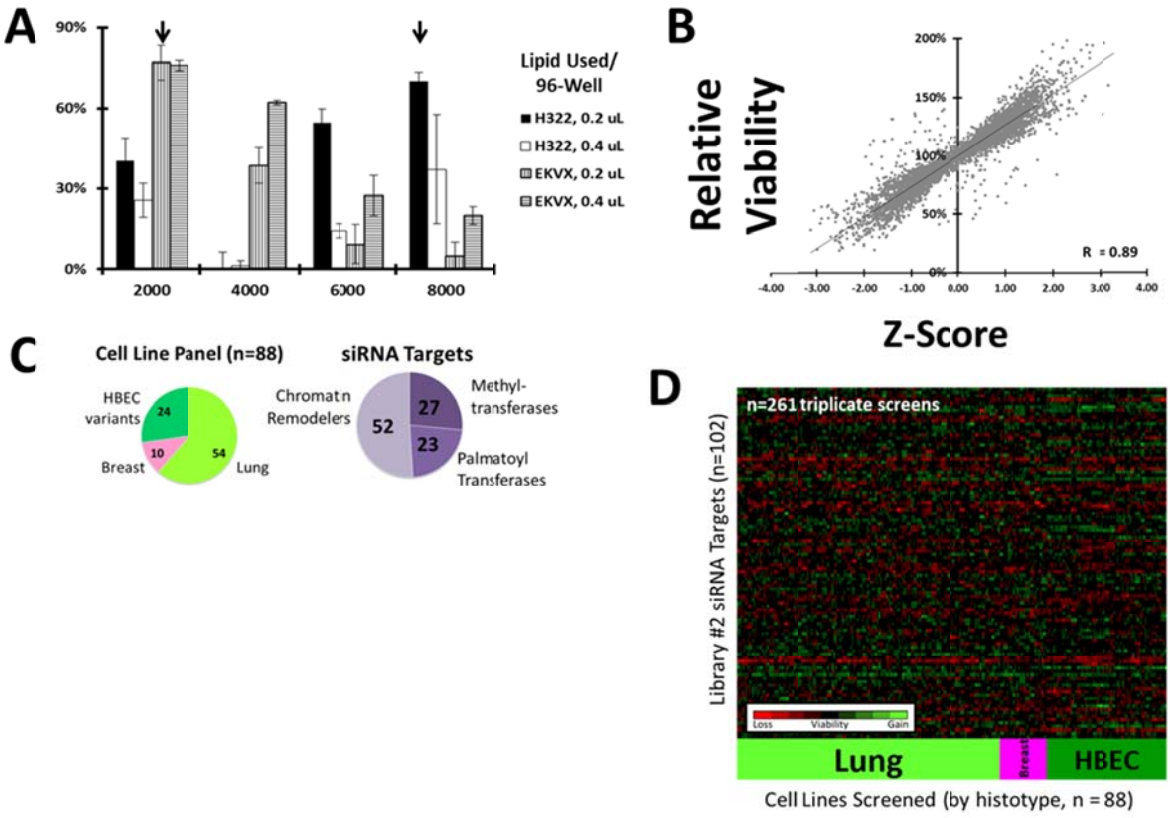


Figure 5.2 Transfection optimization and siRNA screen #2.

A) Representative examples of transfection conditions optimization experiment. Arrows indicate conditions selected for use in screening.
B) Comparison of two different methods of analyzing screening data (viability versus control and z-scores). Pearson correlation.
C) Heat map representation of siRNA screening data for the second siRNA screen (z-scores).
D) Schematic representation of cell lines and siRNA targets interrogated in second siRNA screen.

5.3.2 Screening Dataset (Overview)

Following completion of the screening replicate, data analysis was performed to remove outliers and each screen was checked to see if it clustered with the other replicate screens of the same cell line. In cases where clustering was not seen, additional replicate screens were performed. As discussed in section 2.9, siRNA screening data were processed using two complementary methodologies. Briefly, data were converted to z-scores to allow for direct comparison between different screens. The z-score methodology best accounted for differences in transfection efficiencies as well as reducing the significance of hits in cell lines with many hits and enhancing the significance of hits in cell lines with few hits such that the relative importance of hits could be easily compared. However, this methodology of data processing was found insufficient for complete understanding of the dataset because biological relevance was difficult to assign between like z-scores for different cell lines. For example, a z-score of -1 in one cell line might denote a large reduction in viability while the same score in a different cell line might falsely represent a small reduction in viability as a large viability reduction (presumably due to the small size of the siRNA library leading to large standard deviations in some cell lines with above-average numbers of hits). To address this, the same data were also processed as percent versus plate average (Figure 5.2B). While this methodology does not allow for direct comparison between cell lines, it does allow for easy extraction of biological meaning from the data. Complementary use of both analysis

methodologies in hit identification proved effective in accurately calling specific data points hits or non-hits.

Data were assembled into a heat-map format for broad-scale analysis (Figure 5.1B). While patterns of hits across the panel can be visually inspected to be relatively even between the cancer cell lines of all histotypes, it can be easily appreciated that the HBEC cell lines tested (whether in KSFM media or R5) exhibited a distinct response pattern. Hits could be identified which were toxic to most of the cell lines tested, as well as hits which caused increased viability in the tested cell lines. Unexpectedly, there were relatively few siRNAs tested which exhibited little effect. Furthermore, siRNA knockdown efficiency was also assessed. Thirteen siRNA pools with interesting phenotypes were selected and their knockdown efficiencies were determined in three cell lines (Figure 5.1C). Overall knockdown efficiency was greater than eighty percent with all thirteen siRNAs achieving at least seventy-five percent silencing efficiency.

To assess the relative uniqueness of the response patterns of each cell line to the siRNA panel, correlation coefficients were calculated for all possible pairwise comparisons of the screening replicates (Figure 5.1D). The data were then clustered and converted to a heat map format. As expected, correlation between replicates of the same cell line was found to be quite strong with average correlation coefficients being approximately 0.86. Unexpectedly, the correlation coefficients between replicates that were of different cell lines

were very low (average 0.28) suggesting that each cell line responded uniquely to the siRNA panel. In fact, it was discovered that cell lines of unknown identity could be properly identified by their unique response pattern to the siRNA screen. Additionally, some areas of higher correlation could be identified. For example, it was seen that correlation between the various parental HBECs was higher than background, although the differences between these parental HBEC lines were still surprisingly stark. Additionally, there were a couple of groups of cancer cell lines that exhibited stronger-than-background average correlation coefficients.

5.3.3 A Second siRNA Screen

Following initiation of the Nuclear Receptor/Co-Regulator (NR/CoReg) siRNA screen, a second siRNA screen was initiated that targeted a related set of genes (Figure 5.2C). Specifically, fifty two chromatin remodelers, twenty seven methyl transferases and twenty three palmitoyl transferases were targeted for a total of one hundred and two targets. It has long been appreciated that nuclear receptors act in large complexes consisting of both Co-Regs and chromatin remodelers (reviewed: (Collingwood, Urnov et al. 1999). Additionally, although much less-thoroughly studied, it has been appreciated that many nuclear receptors interact with methyl transferases (reviews: (Kraus and Wong 2002; Lee, Koh et al. 2002; Barrero and Malik 2006)). The palmitoyl transferase targets were included as a follow-up screen to hits from genome-wide siRNA screens performed through a collaboration between

the Minna and White laboratories. The methodology utilized for the second siRNA screen was identical to that used for the NR/CoReg siRNA screen, and results and quality were found to be comparable. A large panel of cell lines was screened including fifty five lung cancer cell lines, eleven breast cancer cell lines, and twenty three HBEC variants. Data analysis was performed as before (Figure 5.2D).

5.3.4 Cluster Analysis of the siRNA Data

As a preliminary analysis tool, clustering methodologies were utilized to establish relationships within the datasets (see methods 2.11). First, it was immediately appreciated that the NR/CoReg siRNA data were capable of distinguishing between HBEC and cancer cell lines (Figure 5.3A). Secondary analysis focused on relationships between the various strains of HBECs and replicates were un-averaged to allow for more detailed studies. Next, clustering of all the HBEC data revealed the siRNA screening data were capable of distinguishing between those HBECs grown in KSFM and those grown in R5 (Figure 5.3B). It has previously been reported that serum adaptation can drastically alter the tumorigenicity, shape, and epithelial properties of HBEC cells (Sato, Larsen et al. 2013). In addition to serum adaptation, various oncogenic variants of parental HBECs have been developed including those with inactivated p53, those with inactivated p53 and inclusion of oncogenic KRAS, and those with inactivated p53, oncogenic KRAS, and inactivation of LKB1. Furthermore, some of these oncogenically progressed variants have been then adapted to serum. Upon cluster analysis it was determined that the siRNA screen was capable of

distinguishing between all of these different variants (Figure 5.3C). Previously in the lab, isogenic soft agar clones of KRAS 30KT, p53/KRAS had been isolated and tested for their tumorigenic properties. Impressively, the 120-target NR/CoReg siRNA screen could accurately distinguish between those clones which were capable of forming tumors in mice and those which were not (Figure 5.3D).

Following these results, it was next examined whether or not the second siRNA screen would also be able to make these same distinctions. Amazingly, a completely different set of siRNAs was also able to perform the same segregations as the NR/CoReg siRNA screen (Figure 5.4A – 5.4D).

Given the amazing resolving power of either siRNA dataset, it was expected that the NR/CoReg siRNA screen would also be able to distinguish between cancer cell lines of differing histotypic origins. Twenty four breast cancer cell lines and seventeen other cancer cell lines of various histotypes were additionally screened using the same methodology as before. Unexpectedly, clustering of all the data showed no distinction between the different histotypes, but instead appeared to be separating the cell lines based on other characteristics (Figure 5.3 E). In direct contrast, examination of HTG qNPA data (Chapter 3) for the same 120 gene targets across a large panel showed exquisite separation of the samples based on their histotypic origins. Based on these data, it was concluded that the functional state of

each cell line (presumably what was being queried by the siRNA panel) was not necessarily directly derived from the cell line's tissue of origin.

Most intriguingly, when the two dendrograms generated from the two siRNA screens were compared, it was found that not only were the HBEC segregations identical, but also that the clusters of cancer cell lines generated were strikingly similar (Figure 5.3F). Clusters were generated using a principal component analysis (PCA) based methodology (to allow for compatibility with the R software) and the two dendrograms were compared using an R package denoted "dendextend." Bootstrapping-based analysis allowed for calculations of p-values for the observed relationships (Figure 5.4 E, F) and the conserved cladistic relationships were deemed highly statically significant.

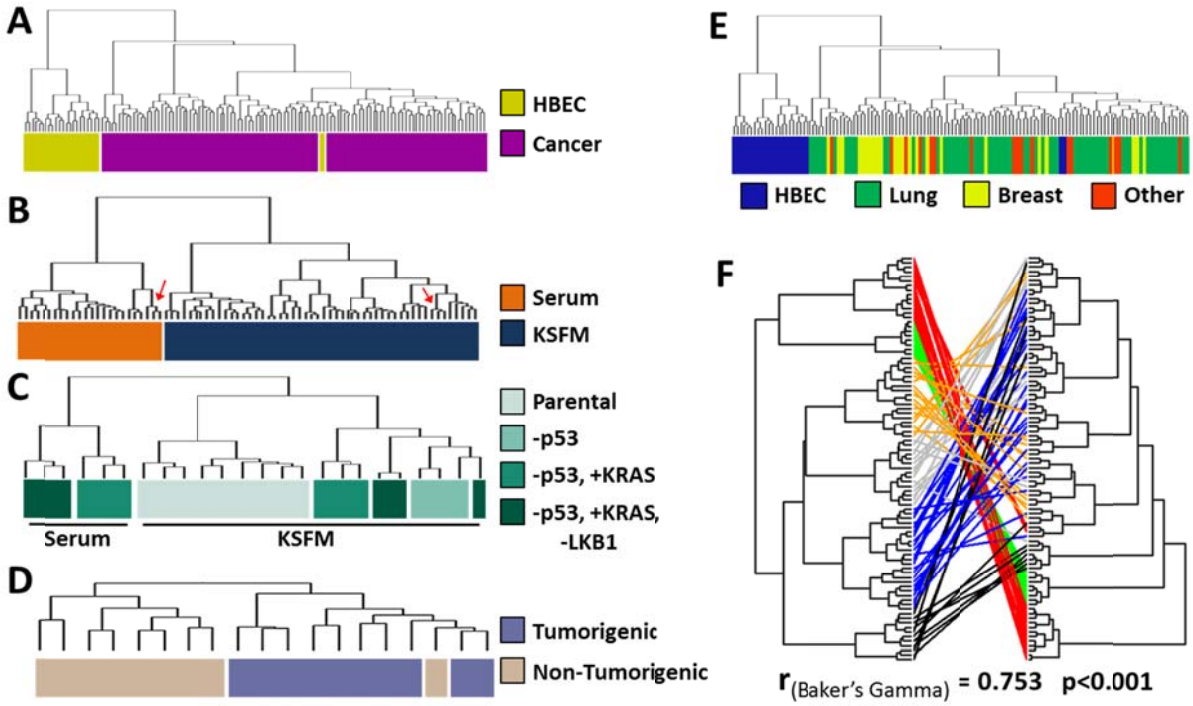


Figure 5.3 Functional relationships within the siRNA dataset.

A) Cluster analysis of all cell lines (n=132). The two serum adapted, oncogenically progressed HBEC variants cluster with the cancer cell lines.
B) HBEC replicate screens (n=76) can be distinguished based on growth media. Red arrows indicate triplicate experiments of HBEC 30KT RL53 in the two different medias.
C) The HBEC 30KT screening replicates can be distinguished based on growth media and the presence/absence of specific oncogenic changes.
D) Screening replicates of the tumorigenic and non-tumorigenic HBEC 3KT RL53 isogenic soft agar clones can be differentiated.
E) Functional distinction cannot be made between cancer cell lines originating from different tissues.
F) A second siRNA screen finds similar functional relationships among the cell lines (n=89).

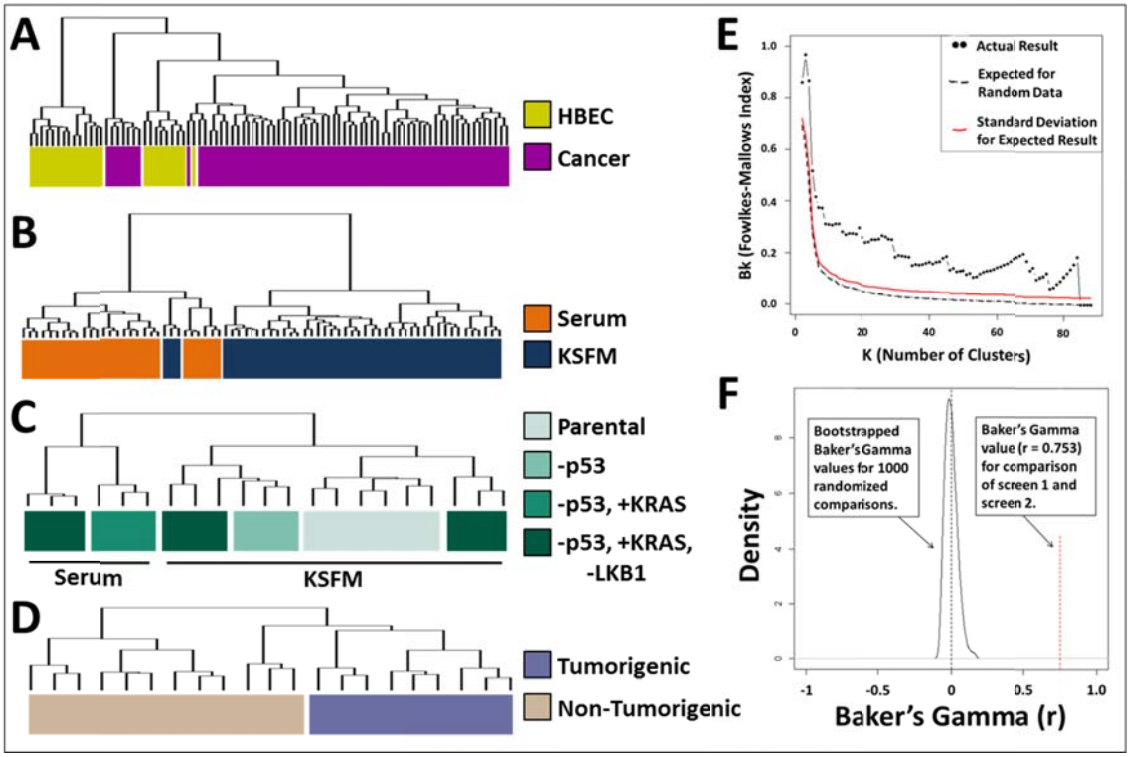


Figure 5.4 Recapitulation of functional relationships in the second siRNA screen.

A-D) Same analysis as in Figure 5.3A-D is shown for siRNA screening dataset #2
E) Fowlkes-Mallows Index calculations for all possible numbers of clusters for siRNA screen #1 dendrogram shown in Figure 5.3F.
F) Result of bootstrapping sampling (N=1000) for siRNA screen #1 dendrogram shown in Figure 4F. Calculated p-value for the similarity demonstrated in figure 5.3F is $p < 0.001$.

5.4.4 Identification of Cancer-Specific Hits

Following the establishment of the quality of the data and the resolving power of the dataset, the challenging process of hit identification and prioritization was begun. As was previously discussed, two complementary methods of data analysis were utilized to best separate true hits from false positives generated due to data analysis methods. To accomplish this, a plot directly comparing results from the two methodologies was generated and analyzed (Figure 5.2B). From the plot, it could be appreciated that any given z-score value included with it an approximate range for biological meaning (i.e. relative kill versus control). For example, a z score of -1 could represent a biological result of 35% loss of viability to 50% loss of viability depending on the cell line context being analyzed. Based on this information, three z-score ranges were selected that best allowed for biological interpretation of the z-score data. Specifically, any result with a z-score greater than -0.8 was assigned a biological result of less than 30% kill and was considered to be "not a hit."

Likewise, z scores between -0.8 and -1.6 were assigned to be between 30% and 40% kill and were considered minor hits. Any z-score that was less than -1.6 had a kill between 40% and up to 90% and was considered a major hit. Hits were then rank ordered inversely on their killing power against the HBEC normal controls and cancer specific hits were then categorized as those which scored no major hits in any of the seven HBEC control lines and two or fewer minor hits in any of the seven HBEC control lines (Figure 5.4A). The majority of hits identified had no major or minor hits in the HBEC control lines.

Most important for clinical translation potential is the identification of a suitable "biomarker" capable of discerning which cancers will be susceptible to treatment. To this end, efforts were undertaken to identify both traditional biomarkers (expression and mutational based) as well as functionally defined biomarkers based on the repeating cladistic relationships derived from the siRNA phenotypes (Figure 5.5B – Figure 5.5D). Figure 5.5B shows an example of an siRNA pool which specifically causes viability loss only in cell lines which are functionally defined as "clade 1" members. Figure 5.5C shows an example of an siRNA pool which specifically kills cell lines exhibiting the functional phenotype of "mesenchymal" (as opposed to "epithelial"). To identify traditional biomarkers, elastic net analyses were performed on cell lines for which both siRNA and mutational/expression data were available. One example of an identified putative biomarker is shown in Figure 5.5D where KEAP1 mutational status appears to predict a kill phenotype for treatment with an siRNA pool.

Given the ability of the siRNA screen to parse between different oncogenic flavors of HBECs (Figure 5.3, 5.4), it is not surprising that there are siRNA hits that have specific effects in virtually every instance of HBEC oncogenic progression. Some examples of siRNA pools that specifically kill based on growth media or oncogenic progression state are shown in Figure 5.5E.

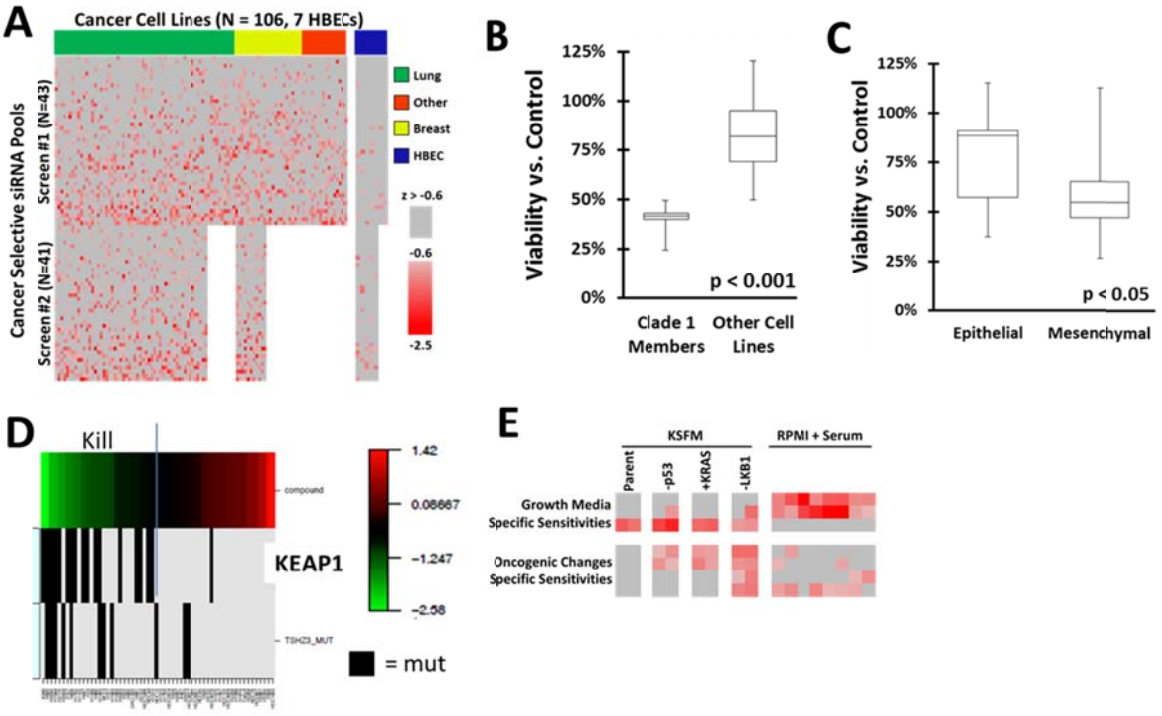


Figure 5.5 Identification of cancer specific hits in the siRNA dataset

A) Hits (total N=84) which cause viability loss in > 10% of the lung cancer cell line panel but exert little effect on the HBEC controls.

B-C) Two examples of functionally-defined biomarkers. In B, the siHMGB2 pool specifically targets members of clade 1. In C, the siSRC3 pool is selectively toxic to mesenchymal cell lines.

D) Elastic-Net analysis of mutational and expression data identifies traditional molecular biomarkers associated with siRNA phenotypes. One example of the siRXRa pool correlating with KEAP1 mutations is shown.

E) siRNA hits are found which: 1) specifically target serum or KSFM grown HBECs, 2) specifically target only oncogenically progressed HBEC lines, or 3) can distinguish between different parental HBEC lines.

5.4.5 esiRNA Validation of siRNA Hits

In order to further validate the identified siRNA hits, a second set of RNAi-based screens was designed based on an independent technology known as endoribonuclease prepared siRNAs (esiRNAs). esiRNA technology utilizes an RNAse III digestion step to "chop up" a target sequence into small fragments that are then used as a pool of siRNA reagents (see methods section 2.14 for details) (Kittler, Surendranath et al. 2007). An initial esiRNA library was produced that targeted the exact same 120 NR/CoReg targets as the siRNA NR/CoReg siRNA library. This library was screened against a panel of twenty-eight NSCLC cell lines, two breast cancer cell lines, and three HBEC cell lines with data processed by the same two complementary methods as before (Figure 5.7A,B). Unexpectedly, the esiRNA screen resulted in very few hits. qRT-PCR analysis was performed to assess the esiRNA reagent's abilities to knock down their targets in comparison to siRNA reagents against the same targets. In ten head-to-head comparisons, it was found that both reagents knocked down their targets to equivalent levels (Figure 5.6B). Likewise, timecourse analysis for one head-to-head comparison found virtually identical rates and efficiencies of knockdown over a four day period (Figure 5.7C). Correlation analyses were performed between the siRNA and esiRNA screens done on the same cell lines with the same genes. While correlation coefficients between replicate screens for siRNA were strong (average 0.88), and correlation coefficients between replicate screens for esiRNA were strong (average 0.83), it was seen that in all cases direct comparison of siRNA and esiRNA screens yielded a low correlation coefficient (average 0.21) (Figure 5.6A).

To further confirm the findings (and the potential hits from the esiRNA screen), the esiRNA library was resynthesized and rescreened across a subset of the thirty cell line panel (Figure 5.7B). Unexpectedly, hits seen in the first synthesis of the library generally did not repeat, and hits found in the second synthesis of the library were generally not seen in the first. Rudimentary investigation of these phenomena strongly implicated improper digestion of select esiRNA reagents in either synthesis of the library as a mediator of a kill-inducing interferon response (Figure 5.7 E-G). Comparison of siRNA and esiRNA results on cell lines screened in both panels revealed many cases where siRNAs had killed a subset of the lung cancer cell lines, but the corresponding esiRNAs had not (Figure 5.6C). As a broad-scale comparison, correlation coefficients were calculated for all pairwise comparisons of all screens and were analyzed (Figure 5.7D). esiRNA screens demonstrate clear separation from their siRNA counterparts.

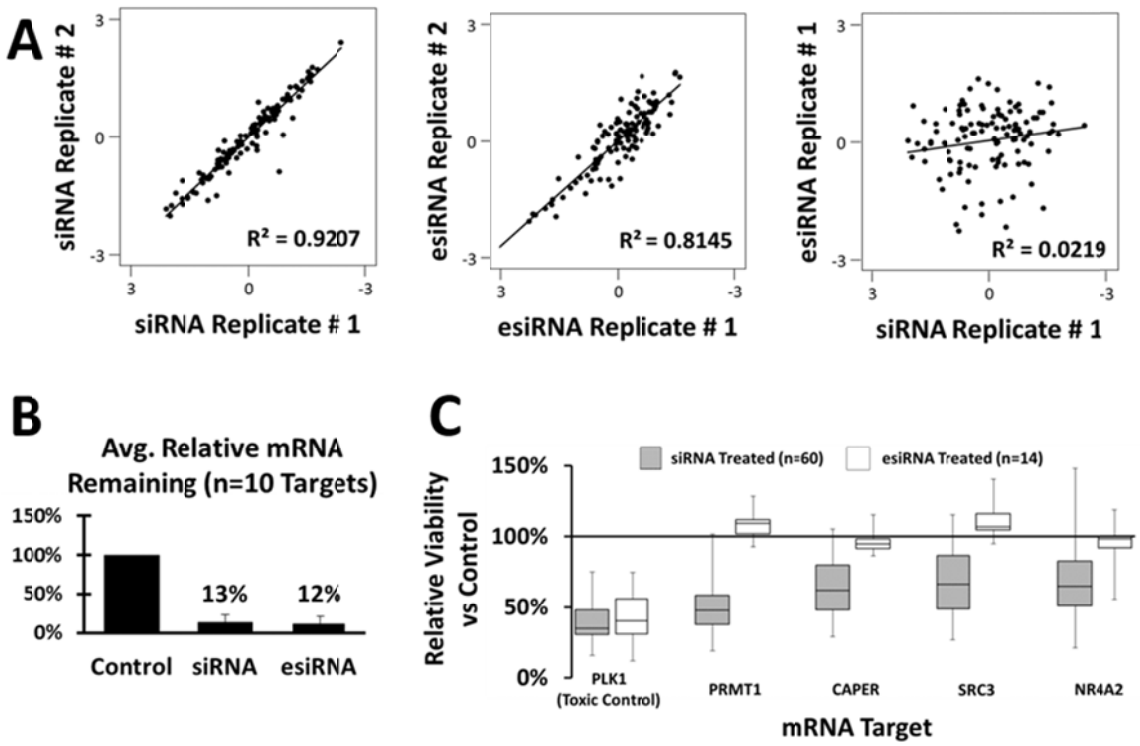


Figure 5.6 siRNA, esiRNA correlation.

A) Representative plots for comparison between results obtained by siRNA or esiRNA screening of the same cell line (Pearson correlation, z-scores, H1993).
B) Average knockdown efficiency of ten siRNAs, esiRNAs across three cell lines (total n = 30, error bars = SD of knockdown efficiency).
C) Box and Whisker plots for selected phenotypic comparisons between siRNA and esiRNA across an identical cell line panel (n=20).

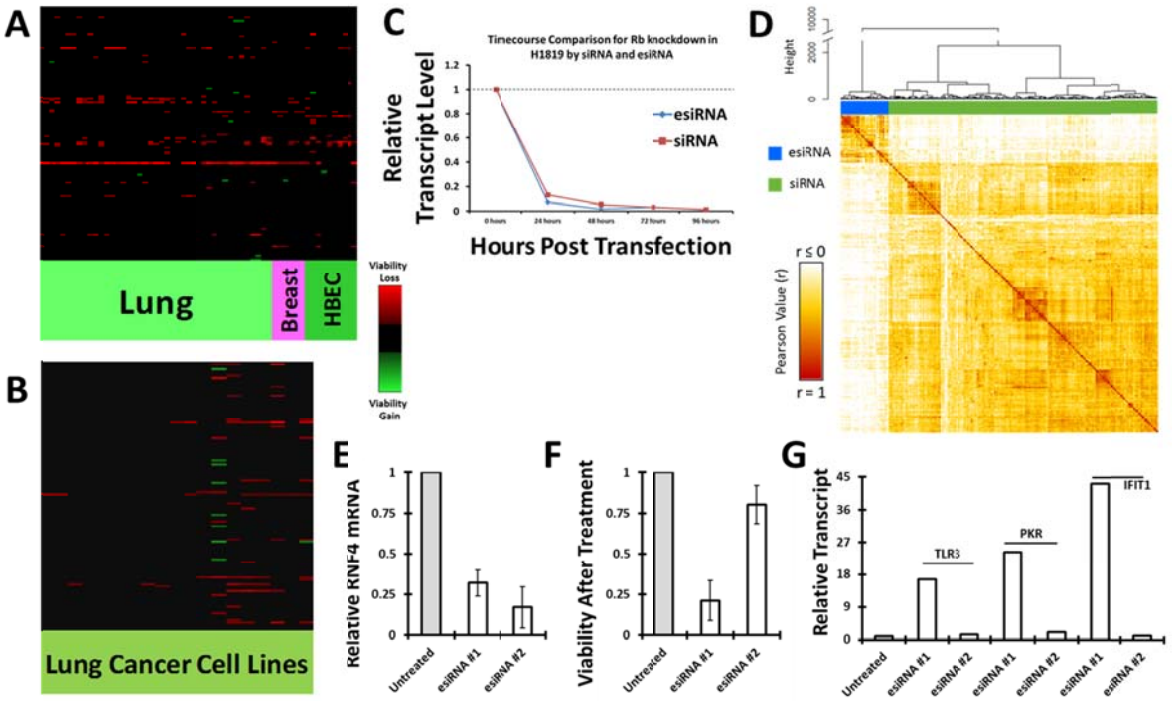


Figure 5.7 esiRNA validation and interferon investigation

A) Heat map representation of resultant phenotypes following screening of two independent the esiRNA library (% versus plate average).
B) Same as (A). Second synthesis of the same library.
C) Timecourse of Rb1 mRNA levels following esiRNA, siRNA knockdown (20nM).
D) Correlation heat map of all screening replicates for siRNA, esiRNA screens.
E) Relative knockdown efficiencies and
F) Subsequent viability changes for two different esiRNAs targeting mRNF4.
G) Relative levels of TLR3, PKR, and IFIT1 (interferon response genes) following treatment with two different esiRNAs targeting mRNF4.

5.4.6 Individual Oligo Analyses

Traditionally in RNAi experiments, off-target effects are detected and prevented by the application of a secondary silencing trigger that produces the same phenotype (Cullen 2006). In the esiRNA experiments, the secondary silencing trigger utilized had not been able to replicate the phenotype of the siRNA reagents initially used. However, to rule out the potential that the incompatible results were the product of the different technological platforms utilized (siRNA versus esiRNA), experiments were undertaken utilizing only the siRNA platform to validate the phenotypes observed in the initial NR/CoReg siRNA screen. Specifically, each pool of four siRNA oligos was rescreened against a small handful of cell lines as unpooled, individual siRNA oligos. If the effects seen with the pooled reagent were due to an on-target effect, it would be expected that each of the individual component oligos of the pool would then yield the same phenotype as the pooled reagent did. However, if the pooled reagent's phenotype were due to an off-target event, the results would not agree.

As expected from the esiRNA results, it was observed that for 119/120 siRNA pools, tests with the individual oligos did not recapitulate the phenotype of the siRNA pool (Figure 5.8A). siRNA pools that caused a cell kill or cell grow phenotype were typically only recapitulated by one or possibly two of the four siRNA oligos comprising the pool. Interestingly, oligos that had no effect in the pooled context sometimes had effects when

used on their own. To rule out the possibility of confounding effects from differing concentrations of oligo, the experiment was repeated at 5nM. While in a few cases the effects were diminished slightly, overall the reduction in concentration did not change the alignment of the individual oligos' phenotypes with the pool's phenotype or the alignment of the 5nM results with the 20nM results (Figure 5.8B, average correlation was 0.774). Furthermore, these results were repeated in 12 additional cell lines (Figure 5.8C).

To further characterize these results, the phenotypes of the four individual oligo treatments were averaged and then compared to the phenotype of the siRNA pools (Figure 5.7B). Interestingly, it was observed that, even in cases where the four individual oligos gave highly disparage phenotypes, the result of the pool could be approximated well by the average of the four individual oligos, suggesting that the pool's phenotype may be somewhat a summation of its parts (Figure 5.8D). Likewise, it was observed that the phenotype of any individual oligo correlated more strongly with the phenotype of the pool of oligos than it did with any other individual oligo (Figure 5.8E).

Despite these results, one pool was identified (siPRMT1) which had its phenotype reasonably well replicated by all four individual oligos. First, these results were repeated and it was verified that the pool and all four individual oligos did in fact give the same phenotype (Figure 5.9A). Next, qRT-PCR analysis was performed and it was determined that the pool and all four individual oligos were highly efficient at knocking down the PRMT1 transcript (Figure 5.8B). Unexpectedly however, further investigations into the mechanism of action of PRMT1-mediated cell killing by cell cycle analysis revealed that each siRNA was killing

through a different mechanism (Figure 5.8E). Oligos #1 and #3 were causing apoptosis while Oligo #2 caused a G1 arrest and Oligo #4 caused a G2/M arrest. Consistent with these results, it was determined that the PRMT1 esiRNA reagent was also capable of knocking down the PRMT1 transcript but that it had no phenotypic effect on the cells (data not shown).

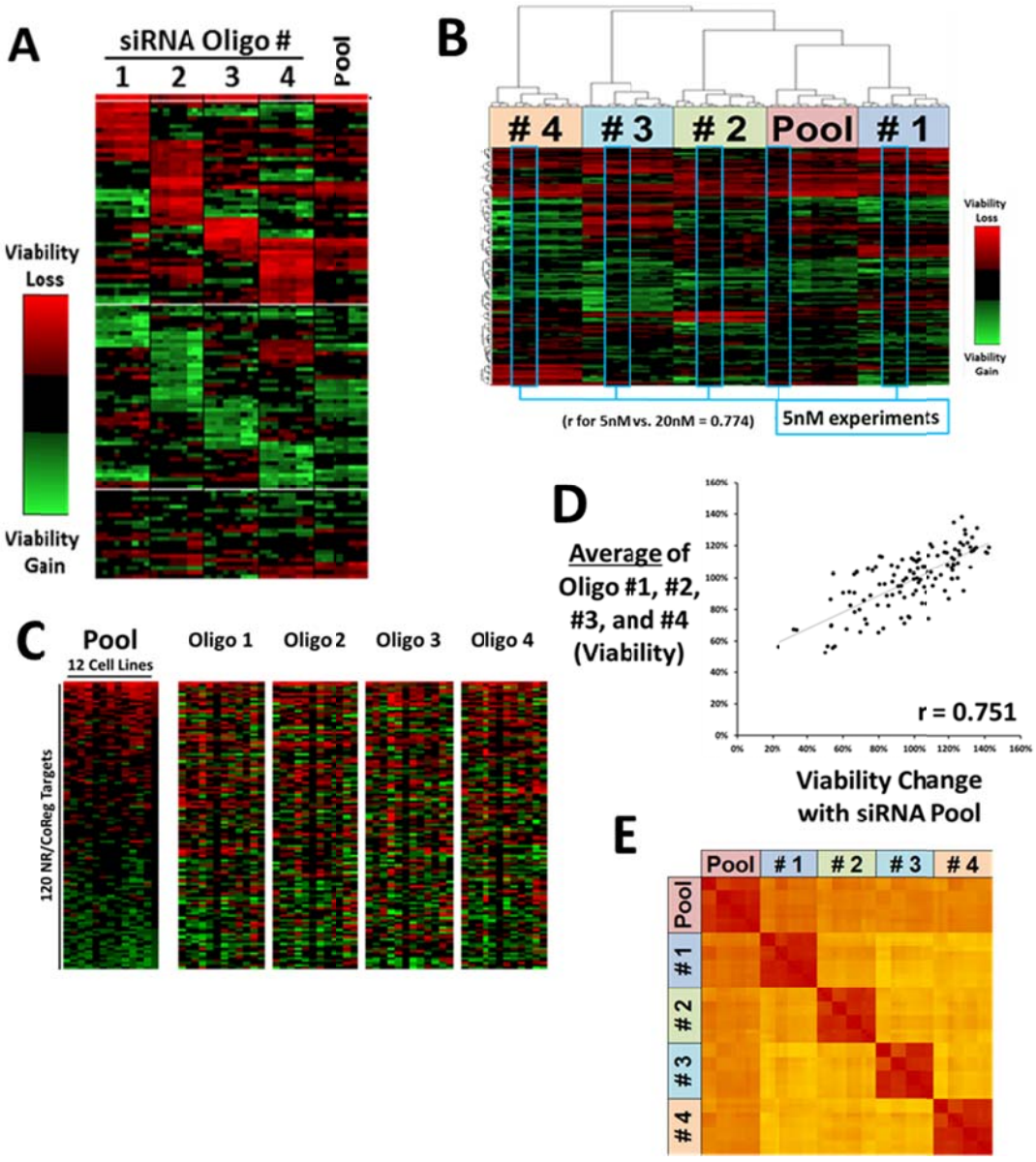


Figure 5.8 Investigation of individual siRNAs as determinates for on-target specificity.

A) The 120 siRNA pools utilized in siRNA screen #1 were broken down into their individual components and rescreened. Results shown as z-scores (H2009).
B) 5nM and 20nM concentrations are shown to be relatively equivalent in the assay.
C) The same assay was done in 12 additional cell lines. Replicates are averaged (n=3).
D) The phenotype of the pool is approximated by the average of the phenotypes of its four individual components.
E) Correlation (Pearson) for each oligo is stronger with the pool than with the other oligos.

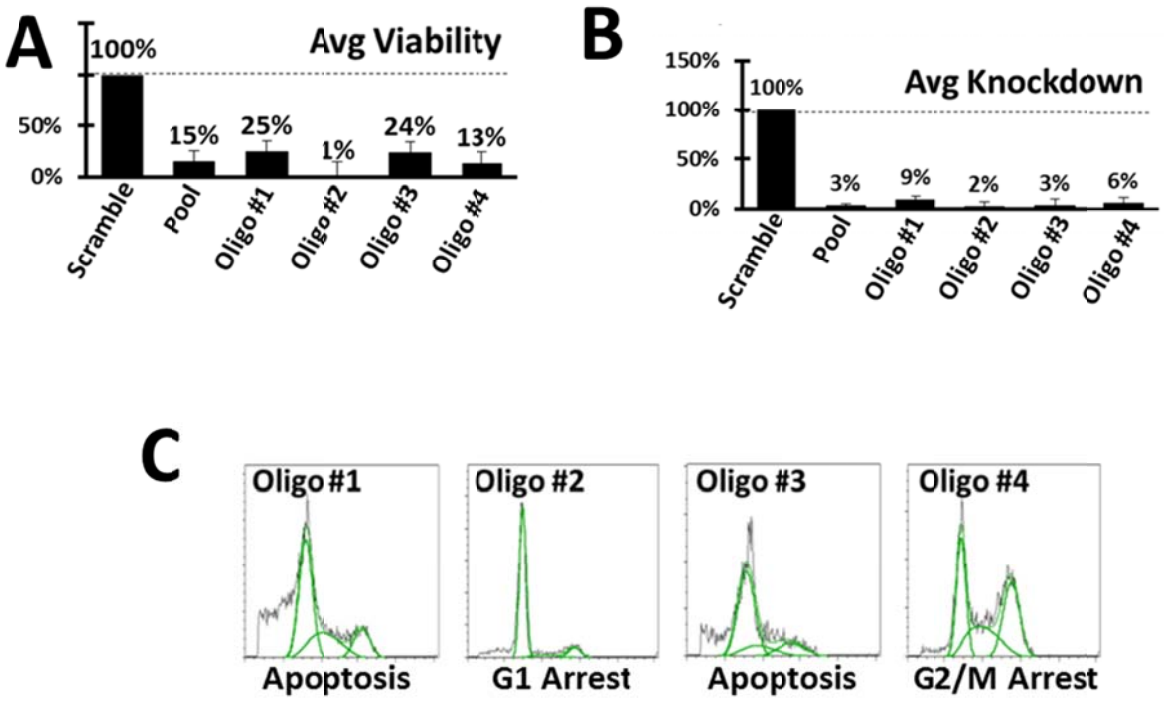


Figure 5.9 Dubious nature of verification by successive oligos.

A) All four siRNAs targeting the PRMT1 transcript are shown to cause loss of viability (average of three cell lines).
B) All four siRNAs targeting the PRMT1 transcript are shown to cause significant knockdown of the target (average of three cell lines).
C) Cell cycle analysis (day three post treatment) for each oligo. Clear differences in mechanism of killing can be observed.

5.4.7 siRNAs Acting as miRNAs

Following these findings, it was necessary to conduct additional experiments to further confirm/characterize the nature of the siRNA off-target effects present in the screening datasets. To accomplish this goal, five siRNA "reagents" were selected for microarray analysis. Two of these pools had specific killing effects against clade 1 (siHNF4a pool and siHMGB2 pool) and two of these pools had specific killing effects against clade 2 (siNURR1 pool and siSIN3B pool). The scrambled control was included as a non-targeting control as well as a non-pooled reagent control. Interestingly, it was determined by western blot and qRT-PCR analysis that HNF4a was not expressed in any of the cell lines tested, necessitating an off-target effect (data not shown).

To assess the unintended silencing effects of treatment with these five siRNA pools, cell lines were treated with lipid only or siRNA reagent, and RNA was harvested twenty-four hours later. Comparison was not made to scramble control treated because (as the data show) the scramble control siRNA was suspected to be responsible for additional off-target effects

and would be a confounding factor in the data analysis. All microarray analyses were performed in triplicate and were analyzed using Matrix Software. Consistent with literature reports, it was seen that treatment of cells with the siRNA pools caused greater than two-fold downregulation of more than the intended on-target transcript (Figure 5.10A). Because no previous literature report had compared off-target signatures between different cell lines, two different cell lines, H157 and H358, were included in the analysis. Interestingly, there were large portions of the off-target signature that were found to be in common between the two cell lines while each cell line additionally had its own cell line-specific component of the off-target signature. As will be discussed later, it is likely that the shared portion of the off-target signature is more robust, but future investigations should determine whether nuances in the miRNA processing machinery between different cell lines could account for these differences in observed off-target signatures. Depending on which specific filters are applied to the dataset (filters utilized in previously published off-target studies differ greatly leaving leeway in deciding on what filter to apply), anywhere from fifty to as many as three hundred transcripts are downregulated upon treatment with an siRNA reagent. Based on work that will be discussed later in the document, it was determined that approximately 130 transcripts were downregulated upon siRNA treatment. Furthermore, it is important to note that overlap between these lists is small – indicating the lists generated are specific for each oligo and not simply a list of transcripts that change regardless of which siRNA might be introduced into the cell (Figure 5.12C).

Although previous literature reports made it evident that the overwhelmingly likely mechanism for these off-target silencing events was miRNA-like transcriptional silencing, an *in silico* analysis was performed to confirm these effects in the NR/CoReg siRNA dataset (example references: (Birmingham, Anderson et al. 2006), (Doench, Petersen et al. 2003)). TargetScan, a miRNA program used to predict targets of endogenous miRNAs (Lewis, Burge et al. 2005), was utilized to predict the targets for each siRNA used in the microarray analysis. An example positive identification is shown in Figure 5.10B whereby TargetScan predicted the scrambled control siRNA would knock down SIX4 mRNA and microarrays before and after siScramble treatment demonstrated SIX4 knockdown by the scrambled control siRNA. For all five of the off-target signatures generated, relative miRNA enrichment was determined (Figure 5.11A). Output from TargetScan is not a yes/no, but instead a scale ranging from 0-100 with 0 being no binding predicted and 100 being binding predicted to be highly likely.

As expected, TargetScan scores were significantly higher for transcripts that were downregulated in both H157 and H358, suggesting that these overlapping portions of the off-target signatures might be more highly enriched for "real" miRNA off-target events. Overall, approximately 79% of the transcripts comprising the off-target signatures were found to consist of transcripts that contained high confidence (TargetScan prediction score greater than fifty) seed sequence binding sites in their untranslated regions (UTRs) (Figure 5.11B). Unexpectedly, it was found that there was a strong enrichment among the transcripts in the

off-target signatures for transcripts containing binding site for more than one siRNA in the siRNA pool (Figure 5.11B). In the case of transcripts downregulated in both H157 and H358, an average of 80% of these transcripts were predicted to be targeted by more than one siRNA from the siRNA pool. This finding suggests that siRNAs within a pooled siRNA reagent may be acting similarly to endogenous miRNAs in that they likely are coordinating together to downregulate their target transcripts. Figure 5.11C shows an example of FAM117B being predicted to be downregulated by all four siRNA components of the siSIN3B siRNA pool.

Because no previous report had examined to what extent the off-target signature of an siRNA pool would be recapitulated by the individual components of that siRNA pool, microarray analyses before and after treatment with each individual component of the siNURR1 pool were undertaken (Figure 5.12A). From the siRNA pool's signature, the following results were observed. Approximately 5% of the siNURR1 pool's off-target signature was not recapitulated by any of the four individual oligos. Approximately 30% of the siNURR1 pool's off-target signature was recapitulated by a single oligo from the siRNA pool (each oligo consisting for approximately 25% of the 30%). An additional 25% of the siNURR1 pool's off-target signature was recapitulated by two different oligos from the siNURR1 pool. Furthermore, approximately 10% of the siNURR1 pool's off-target signature was recapitulated by three of the four individual oligos and the final 5% of the siNURR1 pool's off-target signature was recapitulated by all four of the oligos from the

siNURR1 pool. All the relevant gene lists were analyzed by Ingenuity Pathway Analysis Software (IPA) and it was found that only one out of the one hundred and thirty genes contained in the siNURR1 off-target signature was a known downstream target of NURR1. For each of the individual components of the siNURR1 pool, seed sequence analysis was performed as before utilizing TargetScan to assess the seed sequence enrichment of the off-target signatures generated by these oligos. As expected, off-target signatures for these oligos were strongly enriched (approximately 75%) for high-confidence seed sequence binding sites of the siRNAs used. Interestingly, oligo #2 had a much lower enrichment than the other oligos and furthermore had the weakest off-target signature (Figure 5.12B).

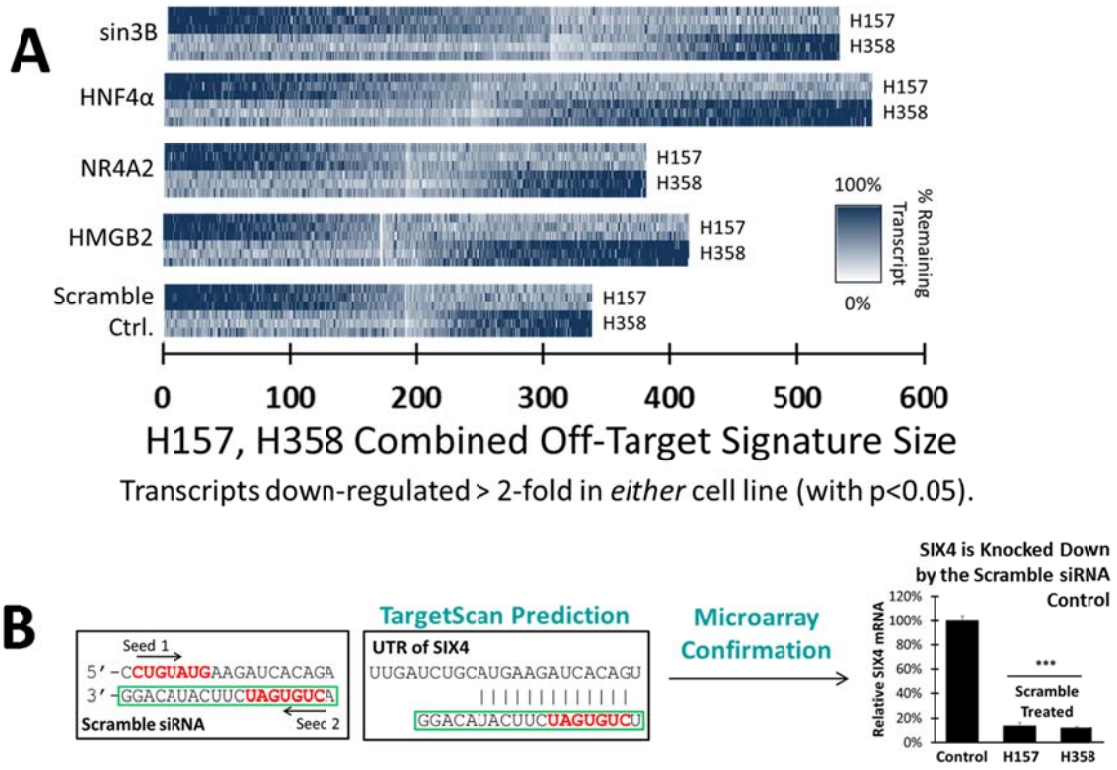


Figure 5.10 Microarray profiling reveals off-target signatures for siRNA pools.

A) Two cell lines (H157 and H358) were profiled by microarray twenty four hours post either mock treatment (lipid only) or treatment with four siRNA pools and the scrambled control siRNA (individual oligo). Triplicate assays are shown un-averaged.

B) An example of a target scan prediction of knockdown of mSIX4 by the scrambled control oligo.

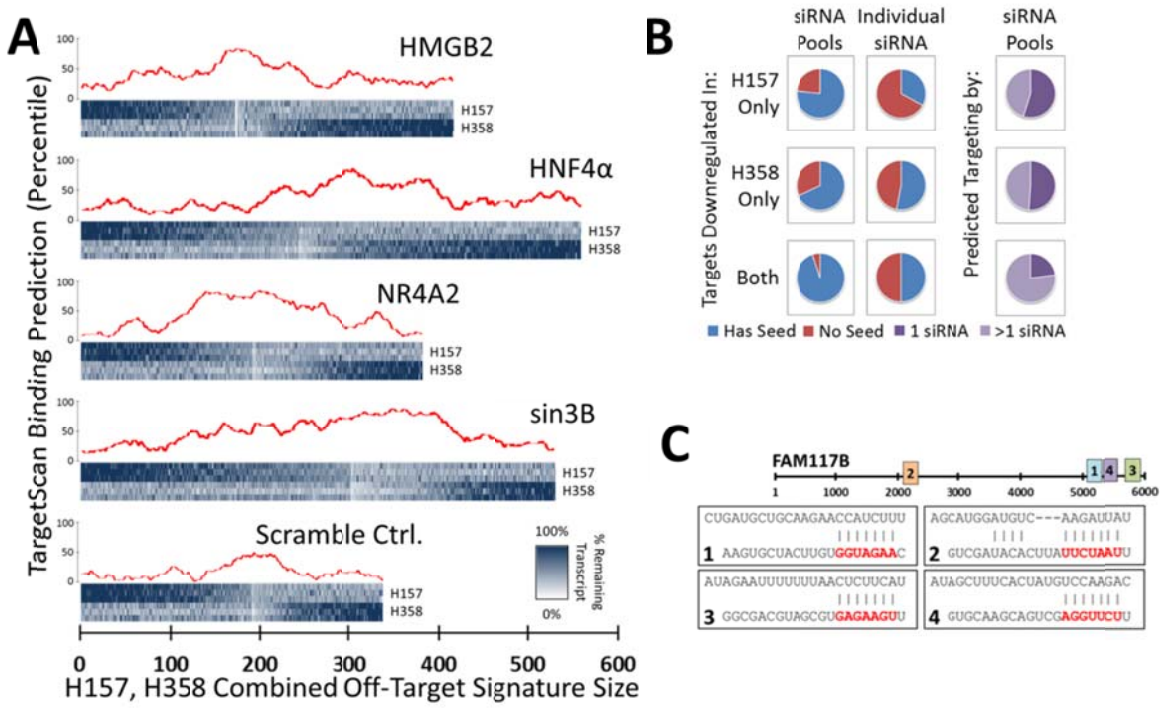


Figure 5.11 Microarray profiling reveals off-target signatures for siRNA pools.

A) Superimposition of TargetScan prediction scores over the off-target signatures. Red line represents a moving average for the prediction score (0-100) where 100 is the strongest prediction. An enrichment for predicted target transcripts can be observed in the section of overlap between the signatures from the two different cell lines.
B) Comparison between signatures and their seed enrichment.
C) Representative example of "multi-targeting" by an siRNA pool. All four components of the siSIN3B siRNA pool are predicted by TargetScan to downregulate FAM117B through a miRNA-like mechanism.

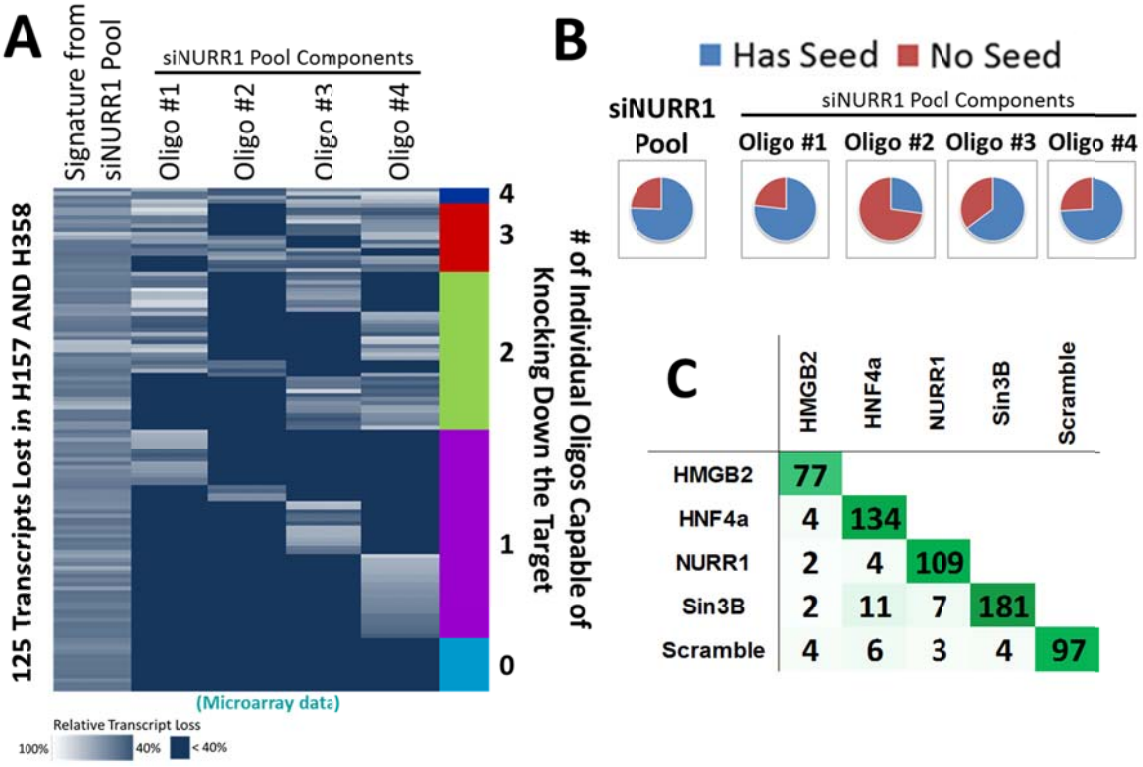


Figure 5.12 Comparison of a pooled off-target signature with off-target signatures generated by the individual components of the pool.

A) Microarrays were run before and after siRNA treatment as in Figure 5.10A. Oligos are sorted as per how many individual siRNA components of the siRNA pool recapitulated the pool's off-target signature.
B) Seed sequence enrichment of the different oligos in the off-target signature.
C) Overlap between the genes downregulated by each siRNA pool/the scrambled control oligo (from Figure 5.10A).

5.4.8 Deconvolution of the siNURR1 Pool: Discovery of N3

Based on the above delineated data, it was necessary to conclude that the majority of the phenotypes observed in the siRNA dataset were caused by miRNA-like off-target effects instead of on-target effects. However, regardless of their origin, the phenotypes recorded in the siRNA NR/CoReg screen were highly reproducible and compelling in terms of their potential for cancer translational opportunities. Specifically, approximately 90 siRNA pools were identified which could kill a large portion of the cancer cell lines but not the normal HBEC controls (Figure 5.5A). Because these phenotypes were compelling, the new goal of discovering the "true target(s)" of an siRNA pool was adopted. Standing at the forefront of this effort, the siNURR1 pool was selected as an excellent candidate for deconvolution analysis because a large portion of the lung cancer cell lines were sensitive to treatment by this oligo (approximately 40% to 60%) while neither the HBEC control lines nor the breast cancer cell lines were significantly affected (Figure 5.13A).

Of first importance, it was necessary to determine which of the four siRNA oligos comprising the siNURR1 pool was responsible for the phenotype or if the phenotype were due to some sort of combinatorial effect generated by the combination of some number of these oligos. Toward this end, variants of each oligo were generated which contained disrupted seed sequences, and all possible pairwise combinations of these functional and

disrupted oligos were tested (Figure 5.13B). It was quickly appreciated that any pairwise combination that contained either a functional oligo #1 or a functional oligo #3 caused cell death while pairs which contained mutated versions of these oligos or only oligos #2 or #4 did not cause loss of cell viability. To confirm this finding, all possible combinations of three oligos (including functional and non-functional variants) were tested (Figure 5.13C). As expected, any combination that contained either a functional oligo #1 or a functional oligo #3 resulted in decreased cell viability while those without functional #1 or #3 did not. Interestingly, combinations containing both #1 and #3 gave a more robust kill than either oligo alone. Presumably this is due to the significant dilution of the individual oligos in the combinations with only #1 or #3.

The finding that either oligo #1 or oligo #3 independently could cause a kill phenotype was unexpected and further experiments were designed to investigate this phenomenon. Specifically, a panel of twenty five cell lines including those most sensitive and most resistant to the siNURR1 pool was selected and each oligo was screened individually across this panel of cell lines (Figure 5.14A). When these results were compared to the results of the siNURR1 pool, it was discovered that treatment of any cell line with oligo #1 caused massive cell death while oligos #2 and #4 caused little loss in viability across the cell line panel. Only oligo #3 (referred to hereafter as N3) was found capable of recapitulating the phenotype of the siRNA pool.

Following these results, a panel of twenty nine cell lines was rescreened with N3 in a six-well dish format (representative results Figure 5.14B). Because transfections were more efficient in 6-well dishes (as determined by comparison of percent kill resultant from toxic treatment in 6-well versus 96-well formats), it was much easier to define sensitive and resistant populations in this format. Additionally, it was discovered that N3 treatment in 6-well format resulted in a binary result (either complete kill or no kill) which was again presumably due to the increased transfection efficiency achieved under these conditions.

Next, efforts were undertaken to attempt to find a biomarker capable of predicting sensitivity or resistance to N3 treatment. Use of Matrix Software on microarray data available for these cell lines found a weak genetic signature with low statistical significance (data not shown). When this signature was run through pathway analysis software GOrilla (Eden, Navon et al. 2009), a statistical association was discovered between the N3 phenotype and upregulation of a mesenchymal phenotype (data not shown). Some previous efforts in the lab had utilized E-cadherin and vimentin western blots to characterize a subset of the cell lines as either epithelial or mesenchymal. Preliminary analysis of these data suggested that N3 sensitive cell lines were statistically enriched for a mesenchymal phenotype. Based on these results, a large-scale effort was undertaken to generate E-cadherin/vimentin western blots for the entire cell line panel represented in the NR/CoReg siRNA dataset (example blot shown in Figure 5.14 C). A complete tabulating of these results is included in Appendix C. Unfortunately, upon expansion of the epithelial/mesenchymal status of the cell line panel, it

was discovered that the initial association of the N3 phenotype with mesenchymal cell lines was no longer statistically significant.

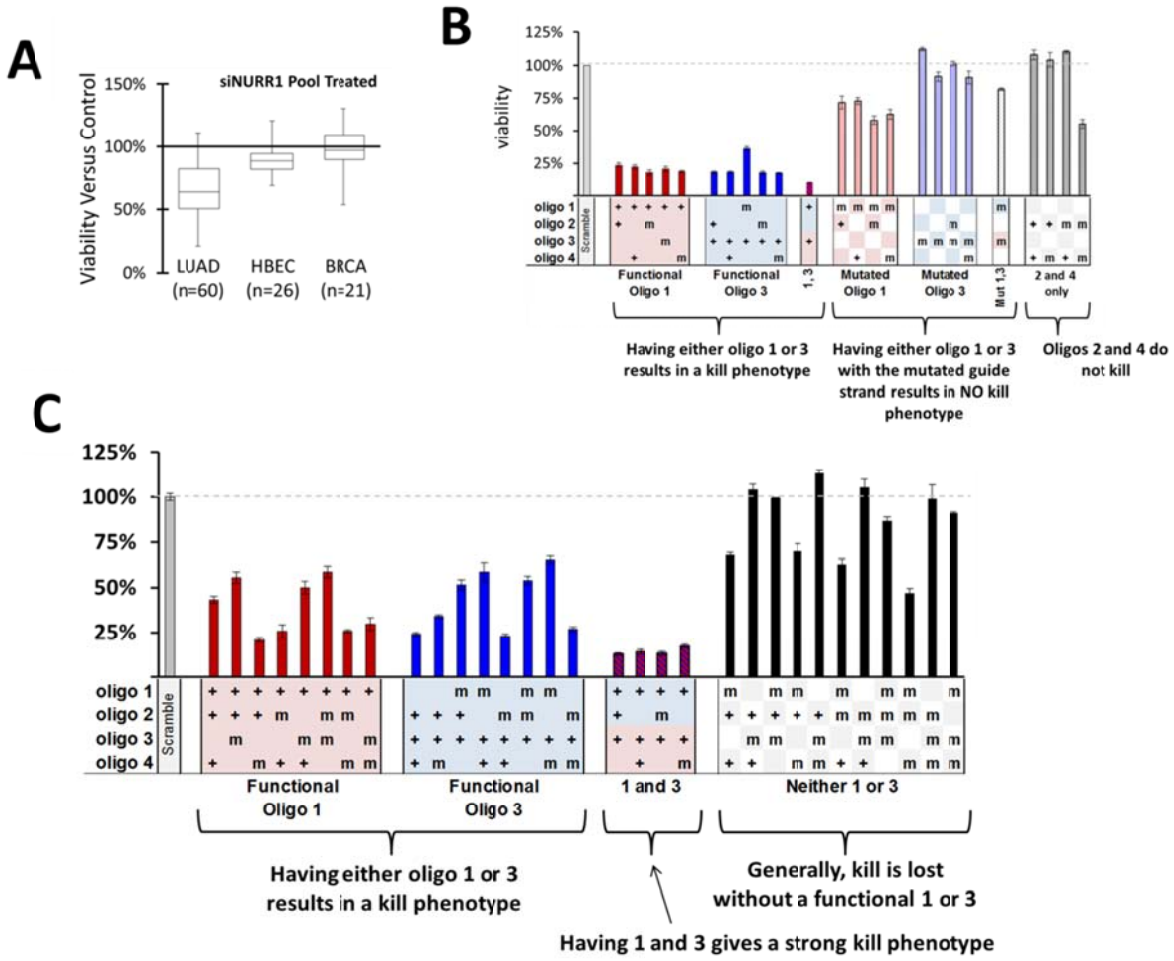


Figure 5.13 The kill from the siNURR1 pool is due to a single oligo--not a combinatorial effect.

A) Box and Whisker plot for the siNURR1 phenotype across the cell line panel. The oligo was found to be more toxic in the lung cancer cell lines than in the HBEC or breast cancer cell lines.
B) All possible pairwise combinations of the four oligo components of the siNURR1 pool as well as mutant version of these oligos (seed region).
C) All possible triplicate combinations of the four oligo components of the siNURR1 pool as well as mutant version of these oligos (seed region).

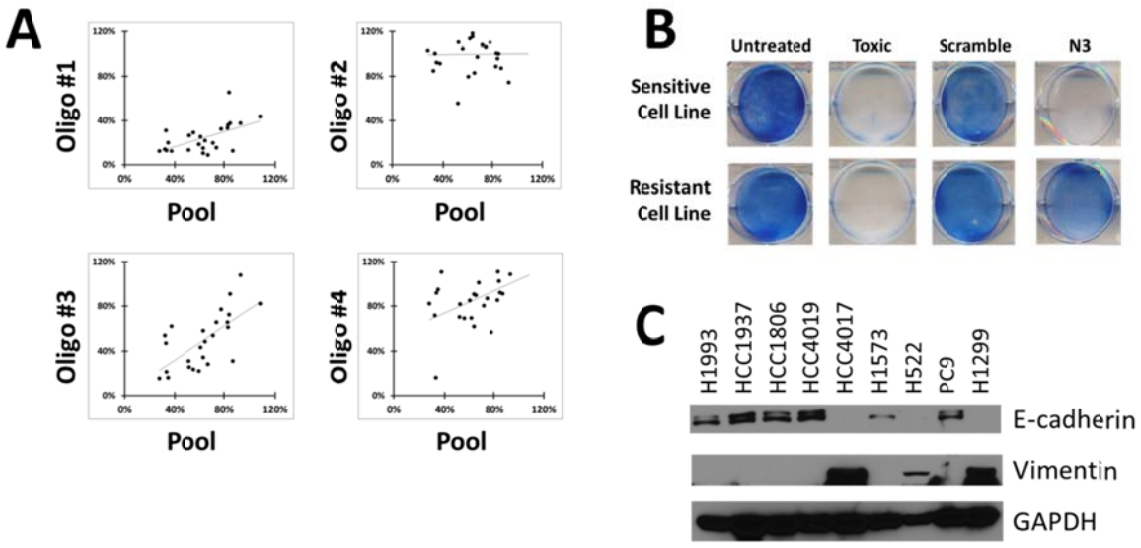


Figure 5.14 Oligo #3 of the siNURR1 pool (N3) is responsible for the kill phenotype of the siNURR1 pool.

A) Each oligo component of the siNURR1 pool was rescreened across a large panel of sensitive and resistant cell lines. Only N3 recapitulated the pool's phenotype.

B) Representative six-well phenotype of a sensitive and resistant cell line five days following N3 treatment.
C) Representative blot for epithelial or mesenchymal (EMT) profiling of the cell line panel.

5.4.9 Mechanism of N3 and Off-Target Confirmation

Next, the mechanism of cell death following treatment with N3 was determined. First, cell cycle analysis was performed on both sensitive and resistant cell lines to examine what effects N3 might be having on the cell cycle (for method details see section 2.15). Based on these analyses it was discovered that N3 treatment induced a robust G2/M arrest in sensitive cell lines but not in resistant cell lines at three days following N3 reverse transfection (Figure 5.15A). Observation of N3 treated wells showed massive cell death between four and five days following reverse transfection. Based on this observation, Annexin IV staining (a test of apoptosis) was performed at four days following N3 treatment. As expected, significant apoptotic fractions could be detected in N3 sensitive but no N3 resistant cell lines (Figure 5.15B). To summarize, N3 treatment caused a G2/M cell cycle arrest three days post reverse transfection followed by an apoptotic event four days post reverse transfection only in sensitive cell lines.

As further confirmation that the phenotype caused by N3 was not due to an on target effect, it was next considered whether or not NURR1 itself might be a biomarker for N3 sensitivity. NURR1 protein and mRNA expression data were assessed for a panel of sixteen cell lines (Figure 5.16A), and no significant association was found between N3 sensitivity and the expression or mRNA level of NURR1. Specificity of the antibody was confirmed by a timecourse of NURR1 knockdown following treatment with N3 (Figure 5.16B). Following these experiments, more sophisticated genetic rescue experiments were designed and performed to confirm whether or not the N3 phenotype was due to an on-target or an off-target effect. To accomplish this goal, two different rescue constructs were designed. The first (dubbed N3-C911) was based on a concept introduced in a 2012 publication (Buehler, Chen et al. 2012). Here it was reported that siRNA on-target effects were dependent on proper target base paring utilizing bases 9, 10, and 11 while miRNA-like off-target effects did not require these binding events. Thus, simple complementation of bases 9 through 11 between the target and guide strands of an siRNA should result in a rescue event in the case that the phenotype is due to an on-target effect.

Analogous to the C911 concept, a second rescue construct was designed that complemented three bases in the seed sequence of the N3 oligo (dubbed N3-Seed). N3 sensitive A549 cells were treated with both rescue constructs and it was observed that only N3-Seed could effect a rescue while N3-C911 elicited no rescue event (Figure 5.17A,B). These results were confirmed by qRT-PCR whereby it was demonstrated that neither N3-

C911 nor N3-Seed could knockdown the NURR1 mRNA transcript (Figure 5.17C). Because neither rescue construct knocked down the NURR1 transcript, but only the N3-Seed construct was able to rescue the phenotype, it was confirmed that the phenotype caused by N3 could not be due to N3's knocking down of the NURR1 transcript. To confirm this effect was not cell line specific, the N3-C911 oligo was tested across the twenty-five cell line member panel and the results were correlated with the N3 phenotype (Figure 5.17D). As expected, treatment with either the N3 oligo or the N3-C911 oligo gave analogous results.

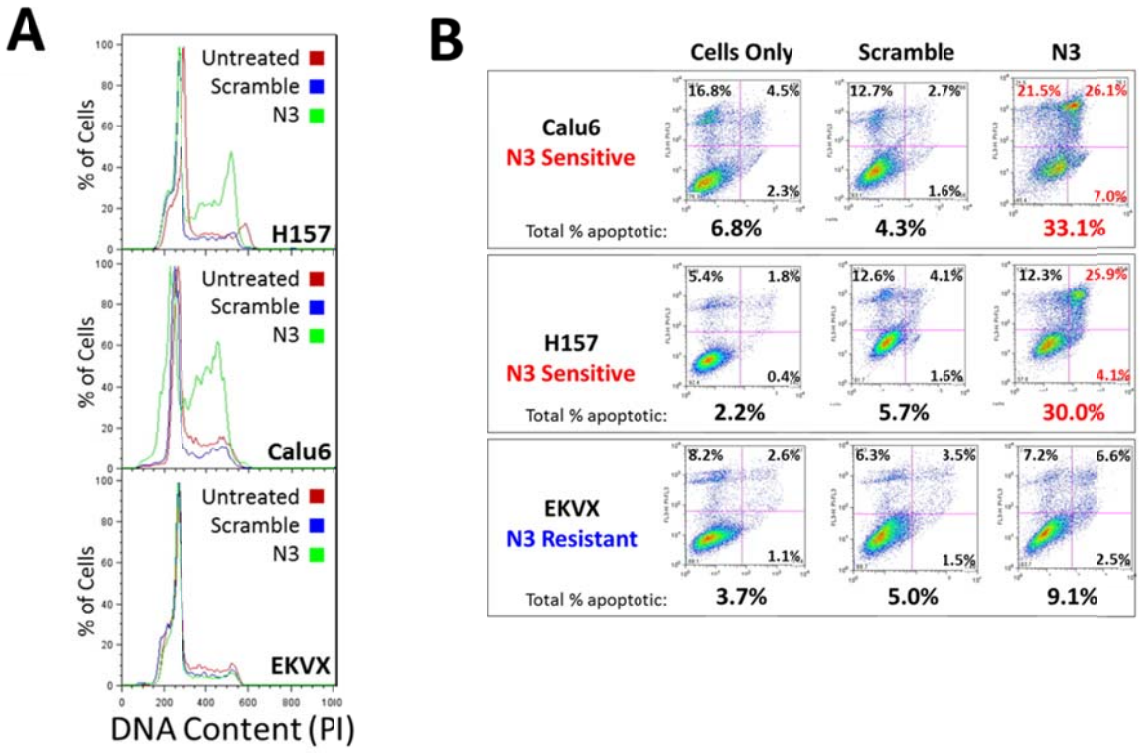


Figure 5.15 Profiling of the N3 kill phenotype.

A) Representative cell cycle profiling of sensitive (H157, Calu6) and resistant (EKVX) cell lines three days following N3 treatment.
B) Representative Annexin V staining plots for two sensitive (H157, Calu6) and one resistant (EKVX) cell lines four days following N3 treatment.

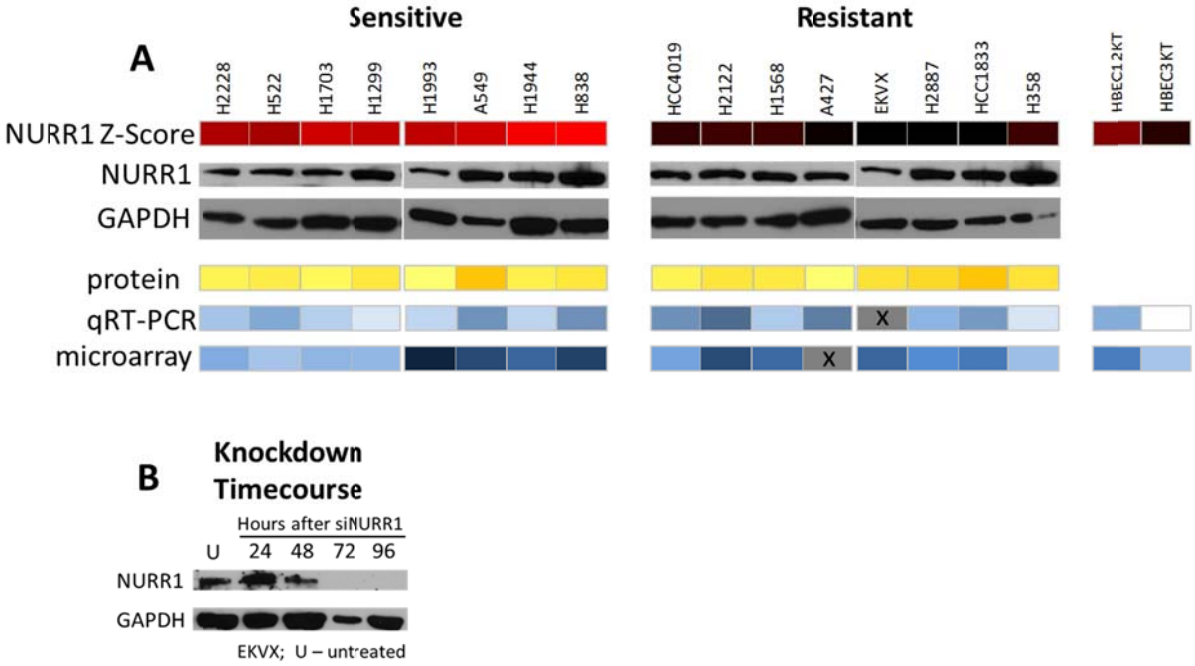


Figure 5.16 Association of the N3 phenotype with molecular biomarkers.

A) No correlation is seen between N3 sensitivity/resistance and the NURR1 protein level or the N3 expression level. Relative protein levels are quantified in the yellow boxes and expression level was assessed both by qRT-PCR and microarray analyses.
B) Time course of N3 knockdown in an N3 resistant cell line (EKVX).

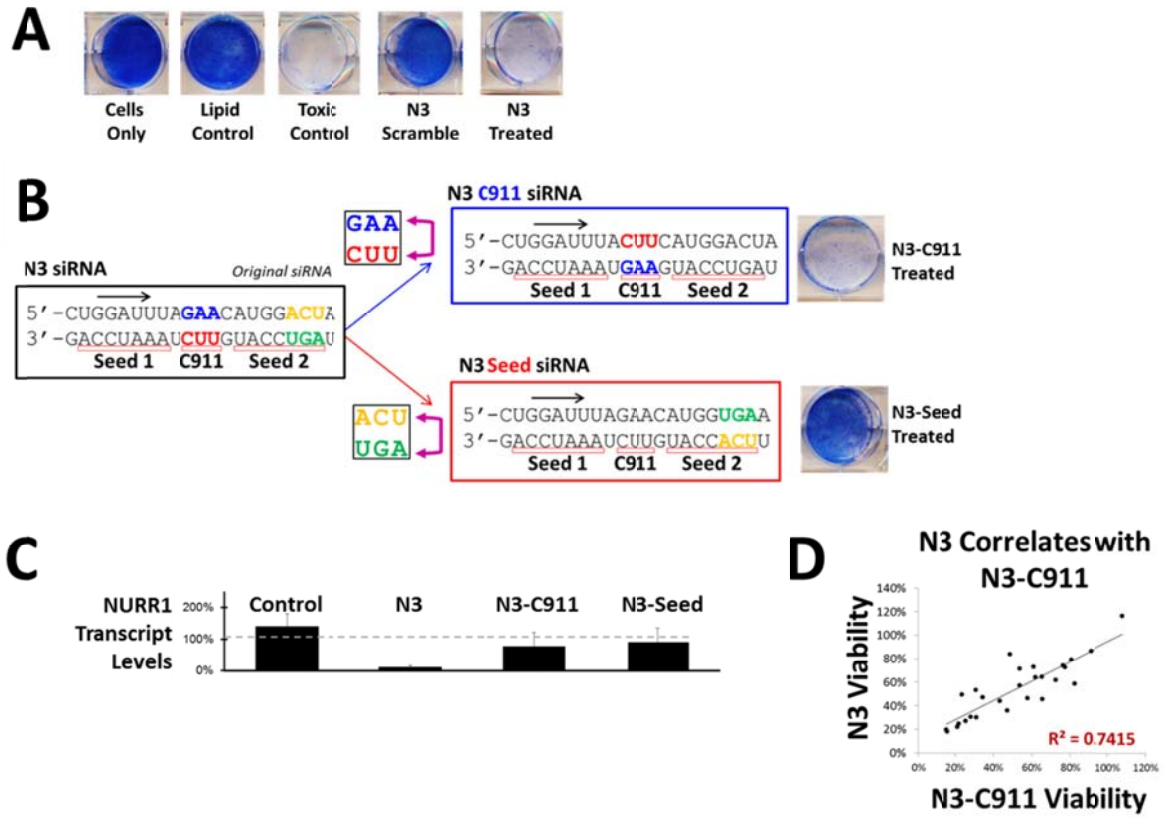


Figure 5.17 The N3 phenotype is driven by a seed sequence effect.

A) Representative images of 6-well dishes stained with crystal violet solution five days following N3 treatment (A549).
B) Schematic representation of the two rescue constructs used and their result (images from same plates as those in 5.17A).
C) Comparison of the ability of the N3, N3-C911, and N3-Seed reagents to knock down the NURR1 transcript.
D) Comparison of the N3 phenotype and the N3-C911 phenotype across the cell line panel (n=25 cell lines).

5.4.10 Identification of eIF4A3 as the True Target of N3

Based on the data gathered to this point, it was conclusive that the N3 phenotype was not mediated through an on-target event. Because the phenotype was so compelling, studies were initiated to discover what the true target of N3 might be. To accomplish this goal, triplicate microarrays were run before and after N3 treatment utilizing four sensitive cell lines (Figure 5.18A). Two possibilities existed as to why the resistant cell lines were not killed by N3 treatment. 1) The resistant cell lines were resistant because of intrinsic molecular differences between the resistant and sensitive cell lines (i.e. some sort of mutational profile or expression difference). 2) The resistant cell lines were resistant because an essential gene was knocked down by N3 in the sensitive cell lines but not the resistant cell lines. Because either possibility could be true, running microarray profiles with resistant cell lines was determined to not be productive as the gene list produced by the resistant cell line(s) would not be helpful in culling the list of genes that still needed to be tested. Following completion of the microarrays, it was determined at a 153 gene signature was reproducibly created in all four sensitive cell lines following N3 treatment. Because no previous literature report had ever compared the off-target signature of a C911 reagent with the off-target signature of the siRNA from which it was derived, we furthermore initiated an experiment to confirm the working nature of the N3-C911 reagent. Importantly, treatment of H157 with the N3-C911 reagent revealed that changing bases 9 through 11 had no discernable effect on the seed sequence caused off-target signature of N3 while still preventing on-target knockdown (Figure 5.18A). These results confirmed that the C911

reagent was working as expected and that C911 reagents are truly capable of recapitulating the off-target signatures of the siRNAs from which they are derived.

Utilizing the information derived from TargetScan previously, the list of one hundred and fifty three transcripts was culled down to ninety eight transcripts by removing any transcript that did not contain a significant binding site for the N3 seed sequence. Other filters utilized include removing genes not significantly expressed in all of the N3 sensitive cell lines or genes which did not have any published literature information. Following the culling of the list, a secondary siRNA screen was performed with a few important changes from the initial siRNA screens. Specifically, because it was observed that siRNA pools were unable to mitigate off-target effects and might even be responsible for enhancing them due to combinatorial off-target effects, a single siRNA was utilized against each target instead of a pool of four siRNAs. Because only a single siRNA was being utilized for each target, a follow-up C911 screen of any positively identified hits was made much more cost effective and simple.

In the secondary screen, four sensitive cell lines and one resistant cell line were screened against the ninety eight target siRNA library. Because the specific pattern of hits produced by N3 was known, all but three of the ninety eight siRNAs could be immediately excluded from the list of potential hits (Figure 5.18B). Importantly, a control siRNA against NURR1 failed to elicit a kill response in any of the N3 sensitive cell lines – again confirming

that NURR1 could not be the cause of the N3 phenotype. An example of a test oligo that did not match the phenotypic pattern of N3 is shown in Figure 5.17B. Three test oligos (eIF4A3, NUDT21, and TM2D2) were found to match the N3 phenotypic pattern reasonably well and these three were further investigated using C911 control oligos to confirm on-target specificity (Figure 5.19A). Using these C911 control oligos, it was quickly discovered that the phenotypes caused by NUDT21 and TM2D2 were due to seed based miRNA-like off-target effects while the phenotype from eIF4A3 was due to an on-target effect as the C911 versions of NUDT21 and TM2D2 were unable to rescue the phenotypes while eIF4A3-C911 gave a complete rescue. To confirm these results, qRT-PCR analysis of eIF4A3 mRNA levels was performed (Figure 5.19B). It was seen that N3, and the siRNA against eIF4A3 both successfully knocked down the eIF4A3 transcript while the C911 version of eIF4A3 was no longer able to reduce eIF4A3 transcript levels.

Several different roles for eIF4A3 have been identified in previous studies. First, it has been appreciated in the literature that eIF4A3 plays a role in the regulation of selenium containing proteins (Budiman, Bubenik et al. 2011). However, the main role in the literature identified for eIF4A3 is as a member of the exon junction complex (EJC) (Chan, Dostie et al. 2004). The EJC can collaborate with UPF1 (Kashima, Yamashita et al. 2006) or CASC3 (Gehring, Lamprinaki et al. 2009) to mediate nonsense-mediated mRNA decay. Furthermore, the EJC has been reported to work with CWC22 to mediate the splicing of exceptionally long introns (Steckelberg, Boehm et al. 2012). Interestingly, a previous screen

done by the White laboratory had identified eIF4A3 and two members of the EJC (MAGOHB and RMB8A) as potentially cancer-specific hits in a similar subset of lung cancer cell lines as those identified sensitive to N3, confirming the likely involvement of the EJC in the phenotype (Kim, Mendiratta et al. 2013) (Figure 5.19C). Furthermore, the same study identified the EJC complex as a specific killer of oncogenically progressed HBECs (Figure 5.19D).

To identify which of eiF4A3's roles might be responsible for the phenotype observed, siRNAs were obtained and tested against N3 sensitive and N3 resistant lines to see which could replicate the N3 pattern across the panel (note: all results were verified using C911 controls (data not shown)). Figure 5.20 depicts a simplified version of the various roles of eIF4A3 and a summary of the results of knocking down EJC cooperative members across the N3 sensitive and N3 resistant cell line panel. Unexpectedly, loss of UPF1 and CASC3 did not replicate the N3 phenotype while loss of CWC22 did – implicating the much more poorly understood role of the EJC in the splicing of exceptionally long introns as the reason for eIF4A3 sensitivity in these cell lines. Unlike the previously published study, loss of MAGOH and MAGOHB did not replicate the phenotype. The most likely explanation is that these two proteins have been reported to compensate for one another in humans and thus loss of either one independently may not have been sufficient to elicit a phenotype (Singh, Wachsmuth et al. 2013).

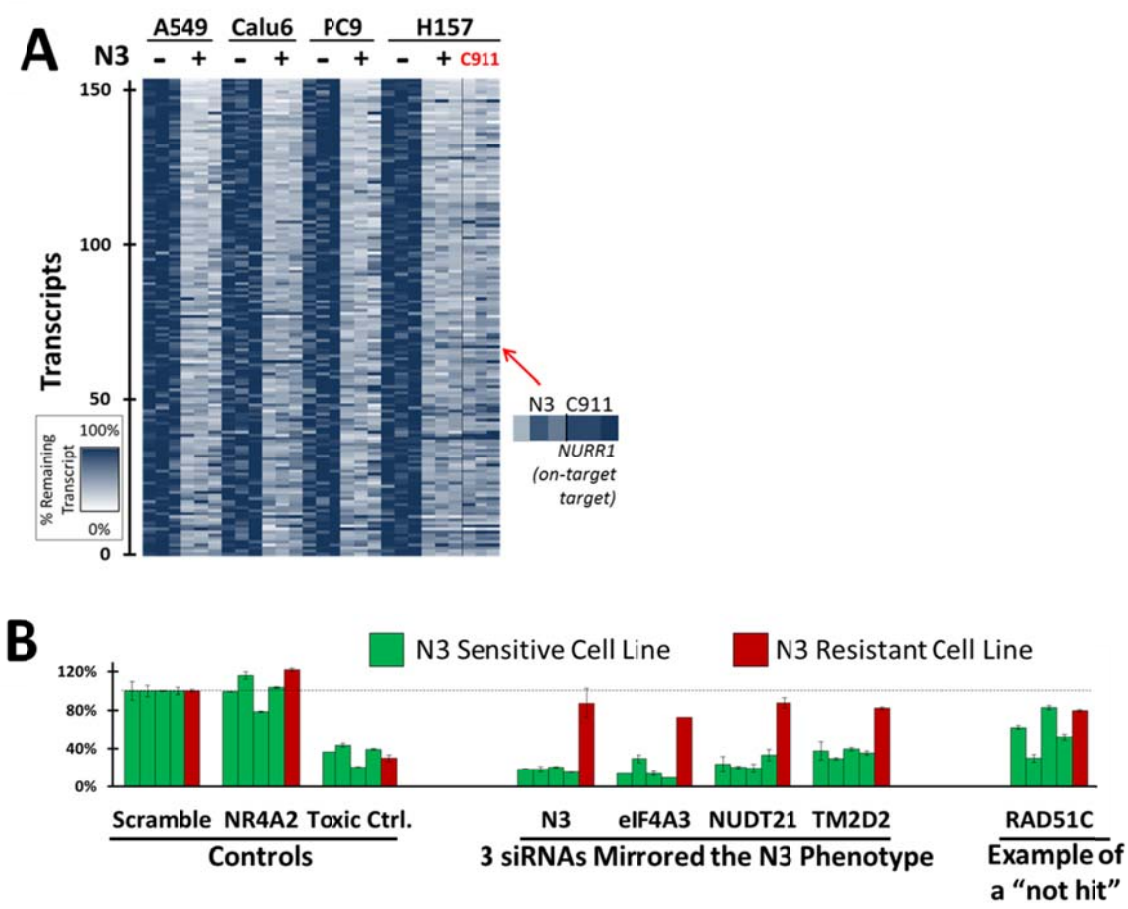


Figure 5.18 N3 off-target signature and subsequent siRNA screening of off-target transcripts.

A) Microarray analyses were run using four N3 sensitive cell lines as before. Additionally, one triplicate set of analyses was run with the N3-C911 reagent to allow for comparison to the N3 off-target signature.

B) Representative results from the siRNA screen. Four sensitive cell lines (same as in panel A) and one resistant cell line (H1819) were profiled against the 100 siRNA panel.

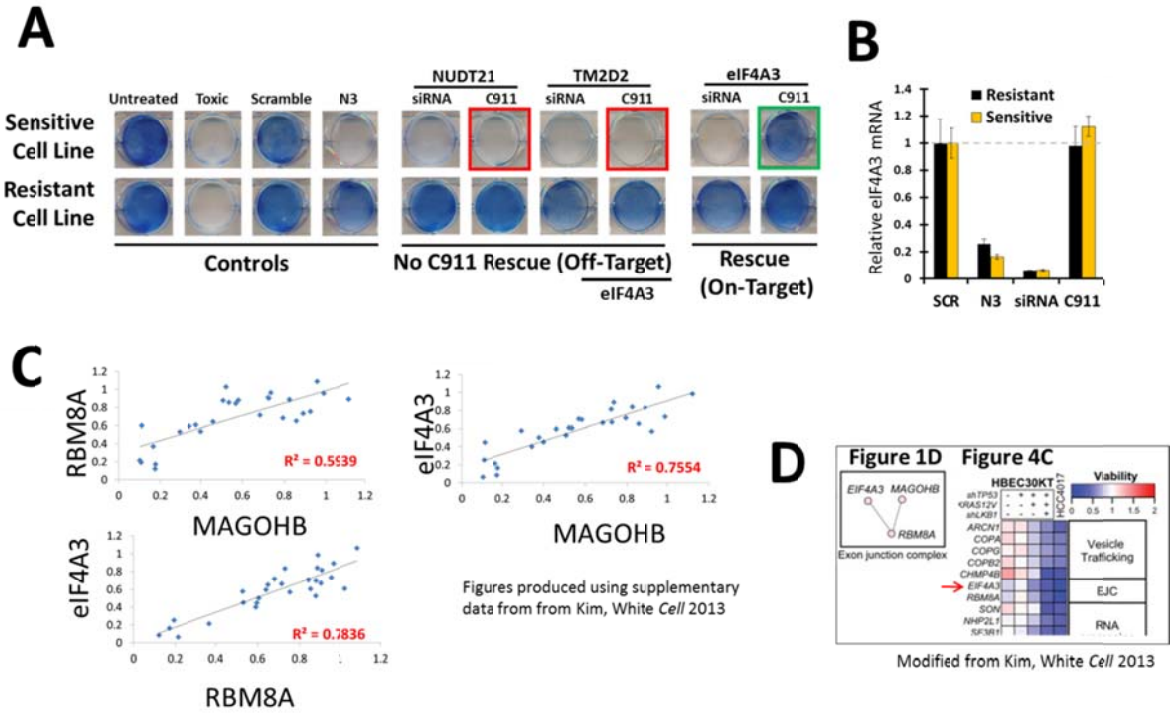


Figure 5.19 Identification of eIF4A3 as the true target of N3.

A) C911 control screening of siRNA hits. Only the eIF4A3-C911 was able to rescue the kill phenotype.
B) Demonstration of knockdown efficiencies of N3, si(eIF4A3) and si(eIF4A3-C911) on the eIF4A3 transcript (qRT-PCR, 24 hours).
C) Modified from Kim, White *Cell* 2013. eIF4A3 and its components were identified as hits in that siRNA screen.

D) Additional figures from the same paper showing the cancer specificity of the eIF4A3 siRNA hit.

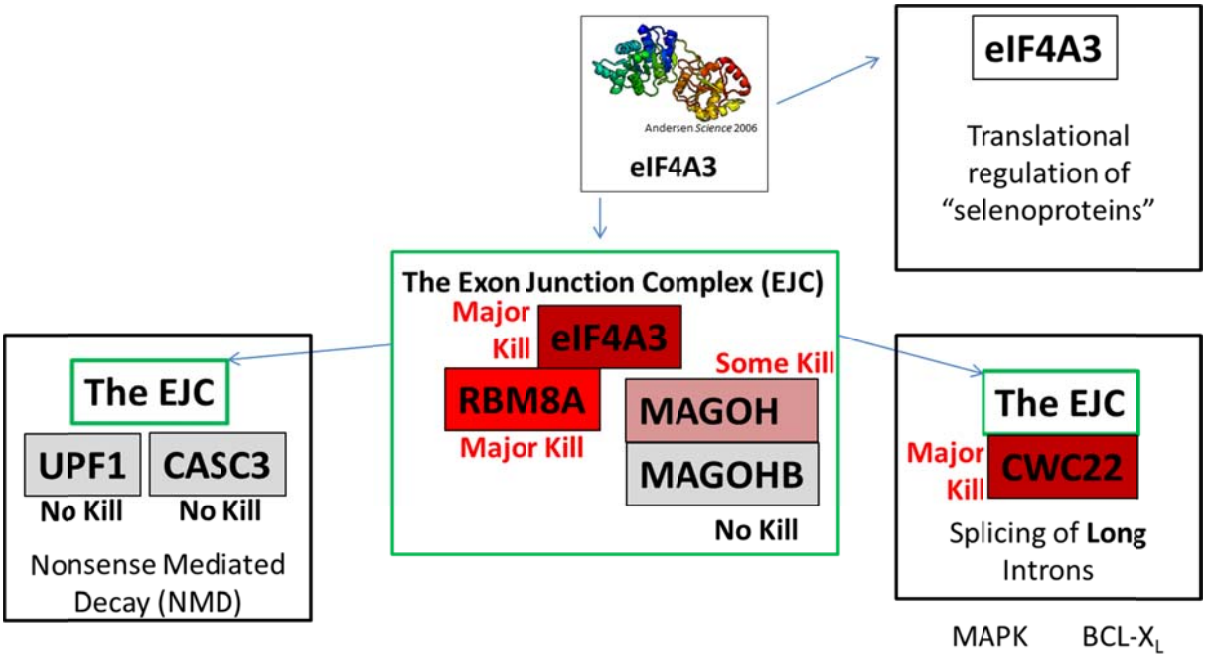


Figure 5.20 Identification of which of eIF4A3's cellular roles mediates the kill phenotype.

Known partners of eIF4A3 were rescreened against a twenty-five cell line panel as individual siRNAs with matched C911 controls. As expected, all components of the exon junction complex recapitulated the N3/eIF4A3 phenotype. Additionally, CWC22, a binding partner with the EJC in eIF4A3's role in splicing long introns recapitulated the phenotype.

5.4.11 eIF4A3 as a Generally Toxic Hit

Following the discovery that treatment of cells with N3 (an siRNA targeted against NURR1) was eliciting a phenotype through knockdown of eIF4A3, confirmation of the cancer-specificity of loss of eIF4A3 was necessary. To accomplish this goal, a panel of twenty-five cell lines was rescreened with N3, siEIF4A3, and a C911 version of siEIF4A3 (Figure 5.21A). Unexpectedly, loss of eIF4A3 was massively toxic to all cell lines tested except for H1819, the cell line used in the follow-up siRNA screen as a representative resistant cell line. In all cases, N3 performed as expected and the C911 version of siEIF4A3 could rescue the kill phenotype confirming that the observed effects were due to an on-target mechanism. While these results might seemingly suggest that eIF4A3 was not the true target of N3, it could not be avoided that N3 did knock down eIF4A3 quite strongly and thus, if eIF4A3 were not the target of N3, no sufficient explanation remained why N3 would not also be completely toxic to all cell lines tested through its knockdown of eIF4A3. Based on these considerations, the hypothesis was developed that cell viability was highly sensitive to even slightly increased or decreased levels of perturbations of mRNA levels of eIF4A3 such that

even small differences in N3-mediated knockdown of eIF4A3 might be sufficient to explain complete cellular kill or rescue.

This hypothesis was tested by simple dilution studies of both N3 and siEIF4A3 in both a sensitive and a resistant cell line (Figure 5.21B). As these experiments are not intuitive, additional explanatory labels were added to the figure. As expected, treatment of Calu6 with N3 caused nearly complete loss of viability while treatment of HBEC 30KT showed little effect. qRT-PCR examination of these treatments revealed an 81% loss of eIF4A3 in Calu6 while HBEC 30KT had only a 76% loss of eIF4A3 transcript. Simple two-fold dilution of N3 was then found to be sufficient to completely rescue the sensitivity of Calu6 to N3. qRT-PCR examination of this treatment revealed that a two-fold dilution of N3 reduced N3-mediated knockdown of eIF4A3 from 81% to 74%, a reduction in efficiency sufficient to rescue the phenotype. Likewise, dilution of siEIF4A3 was observed to follow the same rule. Any treatment which caused more than approximately 77% reduction in eIF4A3 mRNA levels was found to be completely toxic to the cells while treatments which were not sufficient to cause knockdown at this level were found to rescue the phenotype. Based on these results, it was concluded that there existed a razor-thin window for loss of eIF4A3. Unfortunately, the previously identified “cancer-specific therapeutic window” was thus not due to a clinically-tractable reason, but was instead due to slight differences in the transfection efficiencies of various cells.

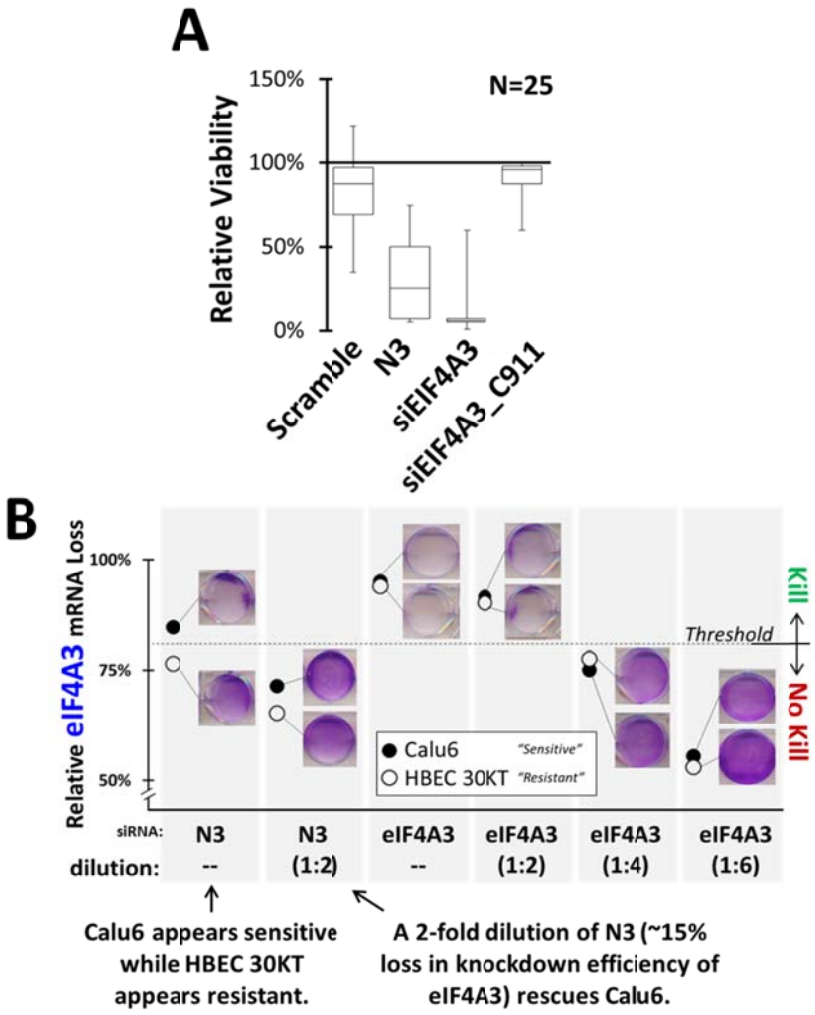


Figure 5.21 Sensitivity/Resistance to N3 is defined by a razor thin threshold of knockdown efficiency of eIF4A3.

A) Rescreening of si(eIF4A3) and si(eIF4A3-C911) across a panel of twenty five cell lines (box and whisker plot, 5 day assay).
B) Relative eIF4A3 transcript level loss following treatment with various concentrations of N3 and si(eIF4A3) in Calu6 (sensitive) and HBEC 30KT (resistant). For any case where knockdown efficiency was greater than ~75%-80%, total loss in viability was observed. However, for all cases with less efficiency than this threshold, little to no kill was observed.

5.4.12 Selection of siRNA Pools for Further Analysis

Following these results, the goal was developed to identify a set of seed sequences that are cancer-specific killers. To accomplish this goal, cancer-specific pools were selected for further analysis utilizing the following criteria:

1. Pick oligos that give maximum coverage across the cell line panel (i.e. every cell line is killed by at least one of the oligos).
2. Pick oligos that have biomarkers.
3. Pick oligos that have activity in the individuals screening that we already did.
4. Pick oligos with potent effects (high magnitude).
5. Pick oligos that are particularly cancer-selective.
6. Try to limit the picks to oligos from the NHR screen.

Additionally, the following technical considerations were taken into account:

1. Need to screen 60, 120, or 180 individual siRNAs (plate layout).
2. Each oligo screened is x3 because of C911 controls and seed sequence controls (i.e. 60 actually =180 siRNAs, 120 actually = 360 siRNAs, 180 actually = 540 siRNAs).

Use of these filters identified 15 siRNA pools that 1) had good magnitude of effect in the screen (>40% kill in >10% of the cell line panel) (Figure 5.22), 2) had >80% coverage of the NSCLC panel (Figure 5.23), 3) were not toxic to the HBEC controls, 4) 13/15 had biomarkers (Figure 5.24; 5/15 had expression only, 4/15 had mutation only, 4/15 had both),

and 5)52/60 of the individual components of these 15 pools were active in the screens we already did.

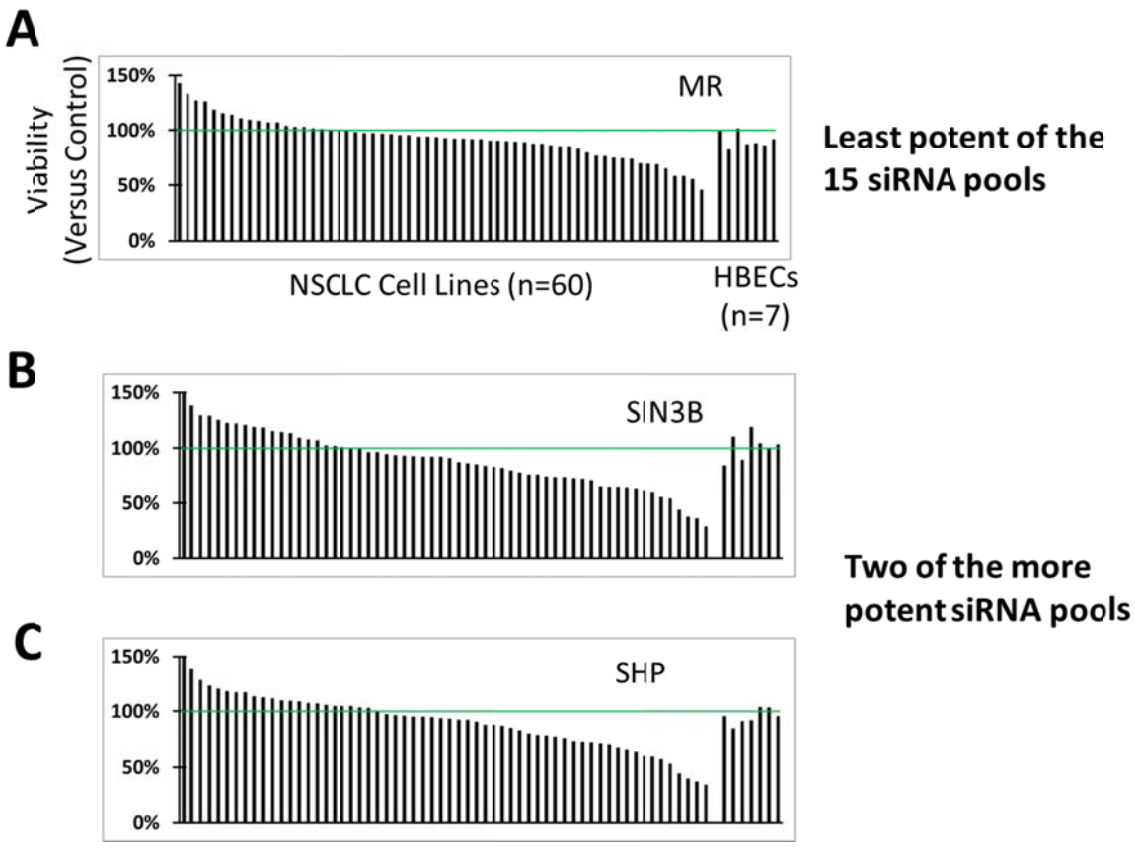


Figure 5.22 Examples of oligos selected.

Oligos were selected for further analysis based on their kill profile across the NSCLC cell line panel. Panels A-C give representative examples of low potency (A) and higher potency (B,C) siRNA pools selected.

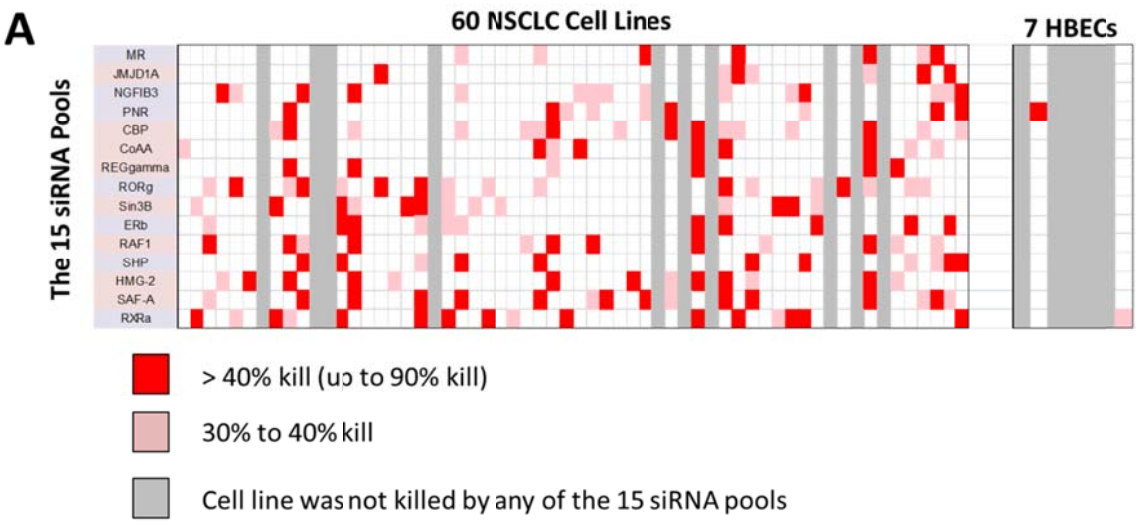


Figure 5.23 Coverage of the Selected siRNA Pools Across the Panel.

A) The 15 siRNA pools selected were found to cause kill in 50/60 NSCLC cell lines analyzed.

A

	expression	mut	
CBP	y		
CoAA	y		
ERb		y	
HMG-2	y		
JMJD1A		y	
MR	y	y	
NGFIB3	y		
PNR			
RAF1		y	
REGgamma			(the 9)
RORg		y	
RXRa	y	y	
SAF-A	y		
SHP	y	y	
Sin3B	y	y	

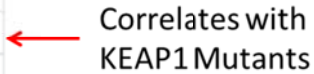


Figure 5.24 Results of the Elastic Net Biomarker Analysis.

A) For each siRNA pool tested, the presence (denoted "y") of either an expression or mutation based biomarker was assessed. The Reggamma siRNA pool did not have a biomarker associated with it, but was found to be associated with one of the functional clades identified previously.

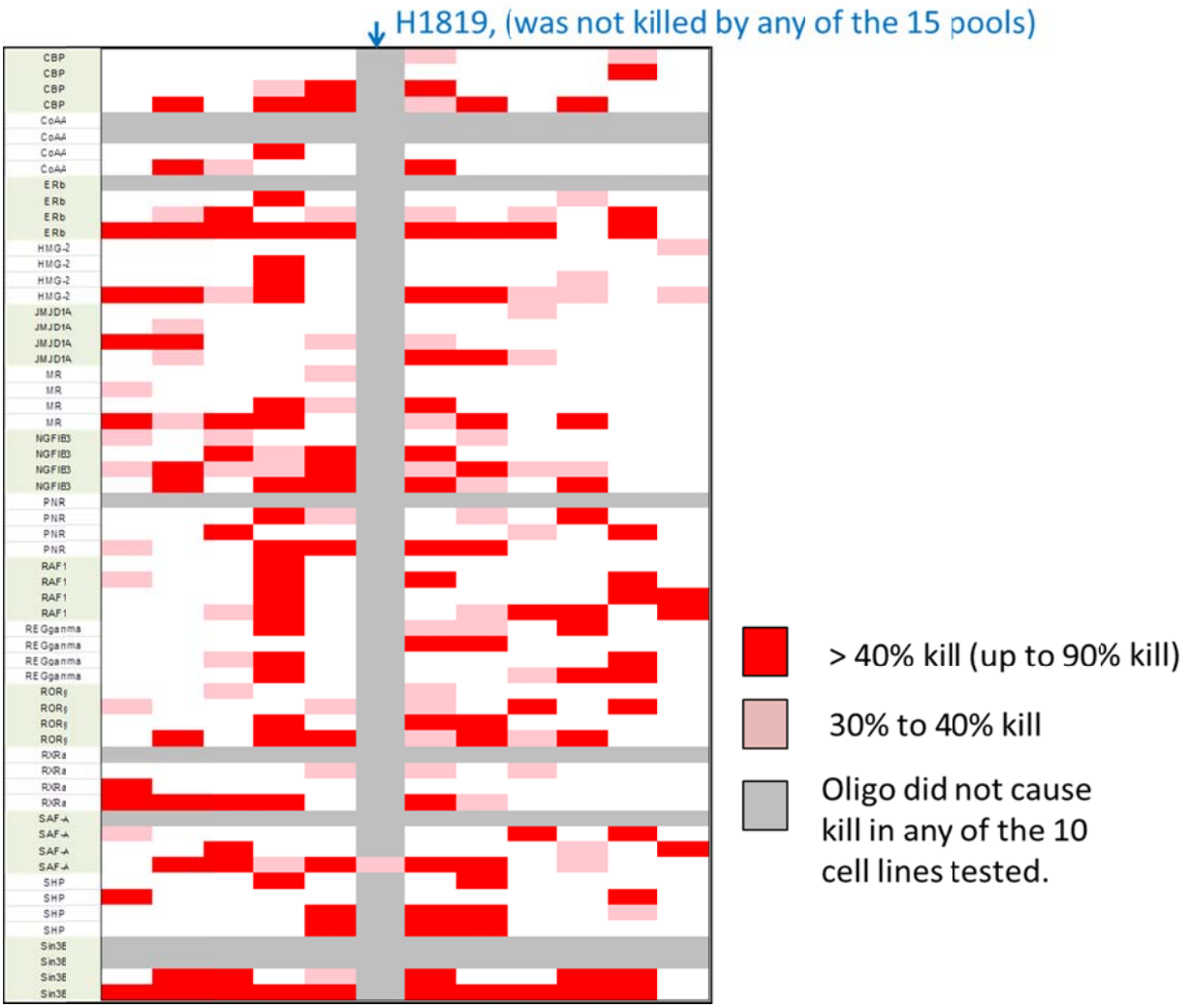


Figure 5.24 Results of the Preliminary Individual siRNA Screens

A) The individual components of these 15 siRNA pools were tested across 10 NSCLC cell lines. Results are shown in the table and are color coded as per the key.

5.4 Discussion

Based on data presented in Chapters 3 and 4, it can be clearly concluded that Nuclear Receptors and their Co-Regulators have a role to play in lung cancer diagnosis, prognosis, and treatment. In direct contrast, data in Chapter 5 (particularly section 5.4.5) would suggest that NRs/CoRegs really have little functional relevance within the lung cancer disease state as their removal by genetic means (esiRNA) produces little-to-no effect across a reasonably sized panel of lung cancer cell lines. These seemingly contradictory pieces of evidence can be reconciled by concluding that NRs/CoRegs are not *drivers* of lung cancer pathobiology (as say KRAS or EGFR mutations might be considered drivers). This observation is supported by the myriad of published datasets which find that NRs/CoRegs are typically not mutated or amplified in lung cancer as would be expected with a *bona fide* driver. Along these same lines, it is furthermore reasonable to conclude that lung cancers are not *addicted* to the continued expression of nuclear receptors or their co-regulators as would be termed in a more classic case of oncogene addiction. Instead, a careful consideration of all the data presented in chapters 3, 4, and 5 leads to the conclusion that, although NRs/CoRegs are not *drivers* in lung cancer, they *are* still important within the disease state (presumably due to

their role as master regulators of key cancer processes), and their manipulation through ligands holds potential for great therapeutic benefit and should be further explored.

Considerations about the role of NRs/CoRegs in lung cancer aside, many important lessons and principles stand to be learned from careful examination of the RNAi screening datasets presented in Chapter 5. First, it can be reasonably concluded that interpretation of siRNA results without thorough consideration of miRNA-like off-target effects is perilous. Evidence supporting this conclusion includes 1) the unexpected resolving power (far beyond what would be expected of the simple phenotypic output of a mere one-hundred genes) of the two complementary siRNA datasets (Sections 5.3.3 and 5.3.4), 2) the complete orthogonal nature of the siRNA and esiRNA screening datasets performed under identical conditions (Section 5.3.5), 3) the microarray datasets and their enrichment for transcripts containing the seed sequences of the siRNAs used (Section 5.3.7), and 4) the inability of the individual components of the siRNA pools to recapitulate the phenotypes observed in from the pool (Section 5.3.6).

Fortunately, most modern RNAi screening pipelines contain steps aimed at removing the most egregious of miRNA-like off-target events. However, reviews of the large-scale success or failure of such methods to actually accomplish the removal of such effects have been mixed at best with many now suggesting these methods to be completely insufficient (Sigoillot and King 2011). More alarming than off-target effects in large-scale screening

efforts where they are systematically tested for and discovered is the potential for RNAi off-target effects to confound everyday bench-level experimental efforts where siRNA reagents are being employed as only a small part of a larger experimental strategy. Specifically, events like those demonstrated in Section 5.4.6 whereby multiple siRNA reagents which all target the same transcript, all knock down the target, and all give similar phenotypes are routinely interpreted as irreproachable on-target events give particular cause for concern.

As a response to these challenges, two avenues of respite are discussed. First, C911 controls offer a simple and cost effective methodology for assessing the presence of an off-target effect. As demonstrated in section 5.4.9, microarray studies with C911 reagents in direction comparison with their non-C911 counterparts find that C911 reagents almost perfectly reproduce the off-target signature of the original reagent while preventing the knockdown of the on-target transcript. Widespread usage of such reagents in both screening protocols (as was demonstrated in Section 5.4.10) and as controls during standard bench-level investigations (as was demonstrated in Section 5.4.11) could revolutionize the RNAi field and allow for continued use of these reagents even in the face of new silencing strategies such as CRISPRi (Larson, Gilbert et al. 2013). In the absence of a true genetic rescue experiment, C911 controls offer a simplistic alternative which can be included in everyday experiments.

Apart from C911 controls, the data in Chapter 5 offer a second solution to handling miRNA-like off-target effects akin to the old adage, "if you can't beat them, join them." Sections 5.4.10 and 5.4.11 demonstrate a novel methodological outline for the deconvolution and identification of the true mediator of a phenotype from an off-target effect. Because genome-wide siRNA screens can be exceedingly expensive to perform, many previous datasets could be mined for hits that were actually mediated by miRNA-like off-target effects and then such deconvolution methodology could be utilized to identify the true targets of these off-target oligos. Such investigations would likely be significantly more cost effective than efforts to repeat such large-scale efforts with alternative technologies. Furthermore, the strikingly similar dendrograms presented in Section 5.3.4 suggest that an oligo library consisting of as few as 120 oligos is capable of saturating virtually all of the functional space within a large cell line panel. Based on this discovery, screening strategies could be easily modified to utilize off-target screens followed by deconvolution steps to allow for screening of large numbers of cell lines in a "multiplexed" fashion.

Finally, as efforts to enable the *in vivo* delivery of siRNA-like complexes reach fruition, it will be important to continue to delve into the potential off-target mechanisms of these reagents. Furthermore, it is feasible to consider reagents designed such that their phenotypes might be affected through miRNA-like mechanisms in addition to traditional on-target silencing to allow for multi-targeted reagents.

5.5 Future Directions

As covered in 5.5 – Discussion, it can be reasonably concluded from the data presented in Chapters 3, 4, and 5 that NRs/CoRegs do not play a role as *drivers* of lung cancer, but that they do perform important functions in terms of regulating general, global cellular processes that may be exploited or unintentionally altered as a consequence of a given lung cancer's pathogenesis. Based on this conclusion, it is unreasonable to continue to pursue the study of NRs/CoRegs as single-agent drivers of lung cancer pathobiology as was the aim of this branch of the study.

However, as is often the case in scientific research, the pursuit of answers about the driver status of NRs/CoRegs has raised additional questions which do merit further study. Of particular import, a better understanding of the true nature of siRNA off-target effects has led to a reinterpretation of the results presented in Section 5.3.2. Particularly, more than one hundred cancer-specific hits can be identified that kill large subsets of the lung cancer cell lines but which do not affect the normal controls or the other lung cancers (Section 5.4.4). Many of these hits were successfully associated with either traditional (molecular) or novel (functional) biomarkers which might be clinically useful for identification of subsets of patients capable of response to these reagents. Two promising avenues for future studies are readily apparent.

First, as the *in vivo* delivery of siRNA-like reagents becomes closer to a reality, the direct application of these reagents to a tumor through such a system becomes a feasible route of investigation. Although in many cases it might prove more advantageous to deliver siRNA-like reagents targeting known driver mutations, many reports exist suggesting that direct targeting approaches may not prove successful (example: KRAS knockdown in KRAS mutant lung cancers only causes cell death in a subset of the mutant liens (Sunaga, Shames et al. 2011)). Indirect targeting through siRNA reagents identified in this study and others like it may prove important for the targeting of these resistant tumors.

Second, Sections 5.3.10 and 5.3.11 present a road map for the deconvolution of an siRNA pool. Efforts have already begun to apply this methodology to additional siRNA pools from the NR/CoReg siRNA screen which appear particularly promising such that their true targets might be identified. Specifically, an siRNA pool targeting HNF4a (which is not expressed in any of the cell lines tested) was found to be a specific killer of a subset of the cell lines (dubbed Clade 1, found in both siRNA screens). Due to its more selective and specific nature, it was considered unlikely that this second siRNA pool was selectively toxic in a similar fashion to the N3 reagent – a hypothesis confirmed by dilution and concentration experiments. Continued investigation of this pool will prove essential in the identification of novel targets which exhibit selective toxicity to defined subsets of the lung cancer cell lines.

APPENDIX A – Transfection Conditions

Cell Line	Cells/well	Lipid/well	Efficiency
A427	4000	0.2	84%
A549	2000	0.4	74%
Calu-6	2000	0.2	83%
DFCI 024	2000	0.2	66%
DFCI 032	2000	0.2	86%
EKVX	2000	0.2	77%
H1155	2000	0.2	67%
H1184	8000	0.2	60%
H1299	2000	0.2	88%
H1355	4000	0.2	67%
H1373	8000	0.2	67%
H1395		inconsistent results	
H1437	2000	0.2	81%
H1568	8000	0.2	65%
H157	2000	0.2	65%
H1573	4000	0.4	65%
H1607	8000	0.2	71%
H1650	8000	0.4	72%
H1693	2000	0.2	81%
H1703	8000	0.2	80%
H1792	2000	0.2	74%
H1819	2000	0.2	69%
H1944	2000	0.2	77%
H1975	8000	0.2	67%
H1993	2000	0.2	66%
H2009	4000	0.2	68%
H2023	4000	0.3	47%
H2030	4000	0.2	71%
H2073	2000	0.2	81%
H2087	8000	0.2	73%
H2122	4000	0.4	67%
H2228	2000	0.2	71%

H226	2000	0.2	69%
H2291	4000	0.4	39%
H23	8000	0.2	83%
H2347	10000	0.4	56%
H2887	8000	0.2	73%
H3122	8000	0.4	80%
H322	10000	0.2	70%
H3255	4000	0.4	56%
H358	2000	0.2	70%
H446	2000	0.2	60%
H460	2000	0.2	84%
H520	2000	0.2	51%
H522	2000	0.2	72%
H647	2000	0.2	63%
H650	8000	0.2	75%
H661	4000	0.4	77%
H820		inconsistent results	
H838	8000	0.4	87%
H841	2000	0.2	80%
H920	2000	0.2	68%
HCC1187	8000	0.2	49%
HCC1438	2000	0.2	75%
HCC15	2000	0.2	64%
HCC1806	2000	0.2	67%
HCC1833	8000	0.2	45%
HCC1954	2000	0.2	79%
HCC2185	8000	0.2	66%
HCC2279	10000	0.2	46%
HCC2405	10000	0.4	58%
HCC2935	8000	0.2	63%
HCC3051	8000	0.2	79%
HCC3153	8000	0.2	78%
HCC4006	4000	0.4	73%
HCC4011	4000	0.4	55%
HCC4017	2000	0.2	68%
HCC4018	8000	0.4	24%
HCC4019	8000	0.2	70%
HCC44	2000	0.2	46%
HCC461	4000	0.2	68%

HCC78	10000	0.2	64%
HCC827	2000	0.2	82%
HCC95	2000	0.2	80%
PC9	2000	0.2	81%
HBEC 3KT	4000	0.3	72%
HBEC 12KT	4000	0.4	58%
HBEC 15KT	2000	0.2	69%
HBEC 30KT	8000	0.2	64%
HBEC 31KT	2000	0.2	59%
HBEC 34KT	4000	0.4	79%
MDA MB 435	4000	0.2	81%
MDA MB 453	4000	0.2	
MDA MB 231	4000	0.2	
MDA MB 468	8000	0.2	79%
MCF7	4000	0.4	46%
SKBR3	4000	0.2	63%
T47D	8000	0.2	72%
BT20	2000	0.2	66%
HCC1937	2000	0.2	63%
HBEC 3KT	4000	0.3	65-75%
3KT-S6Z-L6R	4000	0.3	65-75%
3KT-53	4000	0.3	65-75%
3KT-RL53	4000	0.3	65-75%
3KT-RL53-shL	4000	0.3	65-75%
HBEC 30KT	4000	0.3	65-75%
30KT-S6Z-L6R	4000	0.3	65-75%
30KT-53	4000	0.3	65-75%
30KT-RL53	4000	0.3	65-75%
30KT-RL53-shL	4000	0.3	65-75%
24KT	8000	0.4	74%
HCC1500	8000	0.2	45%
MDA MB 175	2000	0.2	55%
BT 549	4000	0.4	50%
H2405	2000	0.2	66%
HCC1833	2000	0.4	66%
H2347	2000	0.2	63%
H2444	4000	0.2	69%
HCC1500	8000	0.4	48%
HCC515	2000	0.2	64%

HCC1395	2000	0.2	63%
H810	4000	0.2	69%
HCC2302	6000	0.2	50%
ASPC-1	2000	0.2	75%
DU145	2000	0.2	85%
BxPC3	2000	0.2	80%
MIA-PACA2	2000	0.2	80%
PC3	4000	0.2	60%
PANC-1	2000	0.2	60%
H2250	2000	0.2	73%
CAPAN-1	2000	0.2	50%
OVCAR5	2000	0.2	63%
H2258	2000	0.2	76%
LnCAP	2000	0.2	79%
HCC5012	6000	0.2	58%
HT-29	2000	0.4	84%
HCT-116	2000	0.2	45%
HCC1569	8000	0.2	49%
HCC712	4000	0.4	53%
H727	2000	0.2	75%
COLO-205	4000	0.4	14%
SK-MEL-28	2000	0.2	69%
Sk-MEL-5	4000	0.4	42%
SK-MEL-2	4000	0.4	65%
OVCAR4	4000	0.2	65%
KM12	2000	0.4	59%
M14	6000	0.4	55%

APPENDIX B – Ligands and Cell Lines Used

It is important to note that many NR ligands have been reported to interact with multiple receptors depending on what concentrations are being considered. Targets reported in this table are typical targets associated with the listed compounds. Careful consideration of the literature should be undertaken before final assignment of any ligand to any given receptor target(s).

CAS	Brand Name	Product No	Name	Target	MW
152-58-9	SIGMA	R0500	Reichstein's substance S	GR	346.460
4759-48-2	SIGMA	R3255	13-cis-Retinoic acid	RAR	300.440
57-91-0	SIGMA	E8750	alpha-Estradiol	ER	272.380
50-28-2	ALDRICH	250155	beta-Estradiol	Era, Erb	272.380
84852-15-3	ALDRICH	290858	P-NONYLPHENOL	ER	220.350
1806-26-4	ALDRICH	384445	P-OCTYLPHENOL	ER	206.320
103-90-2	ALDRICH	A7302	Acetaminophen	CAR	151.160
52-39-1	SIGMA	A9477	Aldosterone	MR	360.440
302-79-4	SIGMA	R2625	Retinoic acid	RAR	300.440
1224-92-6	sigma	A6401	5alpha-Androstan-3beta-ol	AR	276.460
41859-67-0	SIGMA	B7273	Bezafibrate	PPARa,b,g	361.820
90357-06-5	SIGMA	B9061	Bicalutamide (CDX)	AR	430.370
474-25-9	SIGMA	C9377	CHENODESOXYCHOLIC ACID, PURISS.	FXR	392.570
67-97-0	SIGMA	C9756	Cholecalciferol	VDR	384.640
52214-84-3	AldrichCPR	CRL7135	Ciprofibrate	PPARa	289.150
637-07-0	FLUKA	25895	Clofibrate	PPARa	242.700
23593-75-1	SIGMA	C6019	Clotrimazole	PXR	344.840
50-22-6	SIGMA	C2505	Corticosterone	GR	346.460
53-06-5	SIGMA	C2755	Cortisone	GR	360.440
427-51-0	SIGMA	C3412	Cyproterone acetate	PR / AR	416.940
72-55-9	ALDRICH	123897	1,1-Dichloro-2,2-bis(4-chlorophenyl)ethene	PXR / CAR	318.030
53-43-0	FLUKA	30770	3BETA-HYDROXYANDROST-5-EN-17-ONE	AR	288.420
64-85-7	SIGMA	D6875	21-Hydroxyprogesterone	PPARg, AR	330.460
50-02-2	SIGMA	D2915		MR	
6898-97-1	ALDRICH	218944	Diethylstilbestrol, mixture of cis and trans	ER	268.350
49562-28-9	SIGMA	F6020	Fenofibrate	PPARa	360.830
13311-84-7	SIGMA	F9397	Flutamide	AR	276.210
129453-61-8	SIGMA	I4409	Fulvestrant	ER	606.770
446-72-0	FLUKA	91955	5,7-DIHYDROXY-3-(4-HYDROXY-PHENYL)-CHROMEN-4-ONE	ER	270.240

265129-71-3	SIGMA	G6793	GW7647	PPARa	502.750
22978-25-2	ALDRICH	642355	2-Chloro-5-nitrobenzanilide	PPARg	276.680
50-23-7	SIGMA	H0888	Hydrocortisone	GR	362.460
52806-53-8	SIGMA	H4166	Hydroxyflutamide	AR	292.210
153559-49-0	AldrichCPR	CTG2428	Bexarotene	RXR	348.480
434-13-9	SIGMA	L6250	Lithocholic acid	VDR, PXR	376.570
75330-75-5	SIGMA	M2147	Mevinolin from Aspergillus sp.	PXR	404.540
71-58-9	SIGMA	M1629	Medroxyprogesterone 17-acetate	PR	386.520
21829-25-4	SIGMA	N7634	Nifedipine	PXR	346.330
63612-50-0	SIGMA	N8534	Nilutamide	AR	317.220
112-80-1	ALDRICH	364525	OLEIC ACID	PPARd, g	282.460
57-10-3	ALDRICH	258725	Palmitic acid	PPARa	256.420
57-41-0	SIGMA	D4007	5,5-Diphenylhydantoin	CAR	252.270
50-24-8	SIGMA	P6004	Prednisolone	GR	360.440
1434-54-4	SIGMA	P0543	5-Pregnen-3beta-ol-20-one-16alpha-carbonitrile	PXR	341.490
57-83-0	ALDRICH	850454	Progesterone	PR	314.460
82640-04-8	SIGMA	R1402	Raloxifene hydrochloride	ER serm	510.040
13292-46-1	FLUKA	83907	Rifampicin	PXR	822.940
122320-73-4	AldrichCPR	CTA1760	Rosiglitazone	PPARg	357.430
84371-65-3	SIGMA	M8046	Mifepristone	AR / GR	429.590
52-01-7	SIGMA	S3378	Spironolactone	MR	416.570
10540-29-1	SIGMA	T5648	Tamoxifen	ER	371.510
33069-62-4	AldrichCPR	CSZ9953	Paclitaxel	chemo	853.910
51-24-1	SIGMA	T7650	3,3',5-Triiodothyroacetic acid	THR	621.930
124-94-7	SIGMA	T6376	Triamcinolone	GR	394.430
76-25-5	SIGMA	T6501	Triamcinolone acetonide	GR	434.500
50471-44-8	RIEDEL	45705	Vinclozolin	AR	286.110
17954-98-2	SIGMA	H9384	22(R)-Hydroxycholesterol	PR	402.650
5976-61-4	SIGMA	H4637	4-Hydroxyestradiol	ER	288.380
68047-06-3	SIGMA	H7904	(Z)-4-Hydroxytamoxifen	ER	387.510
102121-60-8	SIGMA	A8843	AM580	RARa	351.440
1153-51-1	SIGMA	A7883	5alpha-Androst-16-en-3alpha-ol	AR	274.440
112965-21-6	SIGMA	C4369	Calcipotriol hydrate	VDR	412.600
87958-67-6	SIGMA	C5749	CGP 52608	ROR	244.340
74772-77-3	SIGMA	C3974	Ciglitizone	PPARg	333.450
338404-52-7	SIGMA	C6240	CITCO	CAR	436.740
2624-43-3	SIGMA	C3490	Cyclofenil	ER	364.430

6217-54-5	SIGMA	D2534	cis-4,7,10,13,16,19-Docosahexaenoic acid	RXR	328.490
10417-94-4	SIGMA	E2011	cis-5,8,11,14,17-Eicosapentaenoic acid		302.450
107724-20-9	SIGMA	E6657		PPARg	
101574-65-6	SIGMA	G6173	GSK 4716	ERRb,g	282.340
317318-84-6	SIGMA	G3295	GW0742	PPARd	471.490
196808249	SIGMA	G5668		ER,AR, PR	
405911-17-3	SIGMA	G6295	GW3965 hydrochloride	LXR	618.510
278779-30-9	SIGMA	G5172	GW4064	FXR	542.840
79558-09-1	SIGMA	L2167	L-165,041	PPARd	402.440
153559-76-3	SIGMA	SML0279	LG100268	RXR	363.490
569-65-3	AldrichCPR	CQP2697	1-[(4-chlorophenyl)(phenyl)methyl]-4-(3-methylbenzyl)piperazine	PXR	390.960
50892-23-4	SIGMA	C7081	WY-14643	PPARa	323.800
126411-39-0	SIGMA	S4194	SR 12813	PXR	504.530
76541-72-5	SIGMA	S1320	SR-202	PPARg	358.650
293754-55-9	SIGMA	T2320	T0901317	LXR	481.330
76150-91-9	SIGMA	T1443	TCPOBOP	CAR	402.060
5630-53-5	SIGMA	T0827			
97322-87-7	SIGMA	T2573	Troglitazone	PPARg	441.540
725247-18-7	SIGMA	X4753	XCT790	ERRa	596.420
362-05-0	AldrichCPR	CPS0490	2-Hydroxyestradiol	ER	288.380
571-20-0	AldrichCPR	CPX4140	ANDROSTANDIOL		292.460
128-23-4	AldrichCPR	CPX9476	PREGNANEDIONE	PR	316.482
63046-09-3	AldrichCPR	CPS7693	6,4'-DIHYDROXYFLAVONE	ER	254.240
459789-99-2	AldrichCPR	FBP1104	6-ETHYLCHENODEOXYCHOLIC ACID	FXR	420.630
170355-78-9	AldrichCPR	EOI6719	4-[7-(1-ADAMANTYL)-6-(2-METHOXYETHOXYMETHOXY)NAPHTHALEN-2-YL]BENZOIC ACID	RAR	486.605
479-13-0	AldrichCPR	CPS6885	Coumestrol	ER	268.220
69552-46-1	AldrichCPR	CQM0824	Carbacyclin	PPARd	350.490
141200-24-0	AldrichCPR	CTB6384	DARGLITAZONE	PPARg	420.487
2265-22-7	AldrichCPR	CQS1760	DEXAMETHASONE 21-METHANESULFONATE	GR	470.555
355129-15-6	AldrichCPR	EGD5847	EPROTIROME	THR	487.142

317318700	AldrichCPR	DYK3000	{2-Methyl-4-{{4-methyl-2-[4-(trifluoromethyl)phenyl]-5-thiazolyl} methylthio}phenoxy}acetic acid	PPARd	453.500
196809-22-0	AldrichCPR	ECA0352	(S)-2-(1-CARBOXY-2-(4-[2-(5-METHYL-2-PHENYLOXAZOL-4-YL)ETHOXY]PHENYL)ETHYLAMINO)BENZOIC ACID METHYL ESTER	PPARg	500.548
247923-29-1	AldrichCPR	DRZ4004	GW-9578	PPARa	492.628
11079-53-1	AldrichCPR	CWV1497	HYPERFORIN		536.792
34816-55-2	AldrichCPR	CQK8990	MOXESTROL	ER	326.433
118414-59-8	AldrichCPR	DPU8322	5-CHLORO-1-[(4-CHLOROPHENYL)METHYL]-3-(PHENYLTHIO)-1H-INDOLE-2-CARBOXYLIC ACID	PPARg	428.337
221368-54-3	AldrichCPR	DFE9999	(R,R)-THC	Era	320.430
76676-34-1	AldrichCPR	CSZ8595	(7A,17A)-17-HYDROXY-3-OXO-7-PROPYLPREGN-4-ENE-21-CARBOXYLIC ACID, POTASSIUM SALT	MR	440.661
151555-47-4	AldrichCPR	CTF3194	RU 58668	ER	658.765
50-53-3	AldrichCPR	CQE3295	CHLORPROMAZINE	CAR	318.870
134404-52-7	AldrichCPR	CTA1599	SEOCALCITOL	VDR	454.690
26538-44-3	AldrichCPR	CSZ4182	ZERANOL	ER	322.398
111025-46-8	AldrichCPR	EBE2136	PIOGLITAZONE-D4	PPARg	360.468
199798-84-0	AldrichCPR	EBY4119	ELOCALCITOL	VDR	442.655
144092-31-9	AldrichCPR	CTF6134	(E)-S,S-DIOXIDE-4-(2-(7-(HEPTYLOXY)-3,4-DIHYDRO-4,4-DIMETHYL-2H-1-BENZOTHIOPYRAN-6-YL)-1-PROPENYL)BENZOIC ACID	RARa	484.653
				GR	

Table of cell lines used in the ligand screen

H1819	HCC1954	HCC202	H2081	H2122	H2347	HCC4006	H1573	H520
H2009	HCC712	BT-474	SKBR3	H2250	MDA-MB-361	HCC4032	H1373	A427
H2073	H1693	HCC1395	HCC2688	H2258	H2052	H1355	30KTshL	H3122

H358	H1993	MDA-MB-231	HME-4077T	HCC827	HCC122	H2882	EKVX	HCC364
H522	MCF7	MDA-MB-453	HME-4064T	MDA-MB-134	HCC2108	H1355	H1944	HCC461
HBEC3-KT	T47D	BT-20	HCC70	H1437	HCC366	H1568	H446	30KT
H157	HCC1806	HCC38	Hs-578-T	H2087	HCC1171	H1792	H1755	30KT_RL53
H1299	ZR-75-1	MDA-MB-157	MDA-MB-361	H2882	HCC3051	HCC1438	H2452	H1838
HCC4017	HCC1569	MDA-MB-468	HCC2302	HCC827	H1650	22RV1	HCC1897	H2172
HCC1833	HCC3153	BT-549	HCC2429	HCC4018	H920	LnCap-AR	HCC2935	H23
H727	HCC1419	HCC1187	H2887	UACC-812	HCC4054	H1975	Calu1	H460
H2170	HCC2185	HCC1143	HCC95	H2126	HCC4058	H2286	H2023	H290
HCC2814	MDA-MB-175	HCC1428	HCC44	HCC15	A549	H650	H2030	H838
BT-483	BT-474	HCC2374	HCC515	HCC193	HCC1359	HCC2352	H2086	
HCC1937	HCC1500	H1607	H1395	HCC78	HCC2450	H1155	H2405	

APPENDIX C – Raw Electronic Datasets

Due to the sheer volume of raw data contained in the datasets summarized herein, it is not feasible to publish these raw values. Additionally, the author realizes that anyone interested in making any use of the data herein will greatly prefer that data in an easily manipulated, electronic format instead of as tables here. Instead, well-annotated electronic files have been created which summarize the data not published here. These files have been placed on the Minna Laboratory Server. The contents of these files are summarized in the table below.

File Name	Description
NR/CoReg siRNA Data	This file contains all the raw data from the siRNA screening efforts processes as both z-scores and "percent-versus control."
Screen 2 siRNA Data	This file contains all the raw data from the second siRNA screening effort processes as both z-scores and "percent-versus control."
esiRNA screening data	This file contains raw data from the esiRNA screens performed processed as percent-versus control and as z-scores.
Ligand screening data	This file contains all the raw IC50 values for each ligand tested across the cell line panel.
HTG Database	This file contains all the raw HTG data used for analyses in this document.
Microarray Data	This folder contains all the raw microarray files (not processed by matrix) used in these analyses.
EMT Database	EMT status of each cell line used in the screen (n=127).

Additionally, a folder of loose files is also included which contains all the excel files used to generate figures in this document. Furthermore, PowerPoint presentation files are also included containing raw versions of the images included herein.

Bibliography

(1999). "A unified nomenclature system for the nuclear receptor superfamily." Cell **97**(2): 161-163.

Abd-Elrahman, K. S., H. M. El-Gowelli, E. I. Saad, A. G. Abdel-Galil and M. M. El-Mas (2010). "Role of PPAR gamma/nitric oxide synthase signaling in the cyclosporine-induced attenuation of endothelium-dependent renovascular vasodilation." J Cardiovasc Pharmacol **56**(2): 195-202.

Altar, C. A., R. A. Hunt, L. W. Jurata, M. J. Webster, E. Derby, P. Gallagher, A. Lemire, J. Brockman and P. Laeng (2008). "Insulin, IGF-1, and muscarinic agonists modulate schizophrenia-associated genes in human neuroblastoma cells." Biol Psychiatry **64**(12): 1077-1087.

Anderson, E. M., A. Birmingham, S. Baskerville, A. Reynolds, E. Maksimova, D. Leake, Y. Fedorov, J. Karpilow and A. Khvorova (2008). "Experimental validation of the importance of seed complement frequency to siRNA specificity." RNA **14**(5): 853-861.

Annibaldi, A. and C. Widmann (2010). "Glucose metabolism in cancer cells." Curr Opin Clin Nutr Metab Care **13**(4): 466-470.

Anzick, S. L., J. Kononen, R. L. Walker, D. O. Azorsa, M. M. Tanner, X. Y. Guan, G. Sauter, O. P. Kallioniemi, J. M. Trent and P. S. Meltzer (1997). "AIB1, a steroid

receptor coactivator amplified in breast and ovarian cancer." Science **277**(5328): 965-968.

Aranda, A. and A. Pascual (2001). "Nuclear hormone receptors and gene expression." Physiol Rev **81**(3): 1269-1304.

Ashkenazi, A. (2002). "Targeting death and decoy receptors of the tumour-necrosis factor superfamily." Nat Rev Cancer **2**(6): 420-430.

Barbie, D. A., P. Tamayo, J. S. Boehm, S. Y. Kim, S. E. Moody, I. F. Dunn, A. C. Schinzel, P. Sandy, E. Meylan, C. Scholl, S. Frohling, E. M. Chan, M. L. Sos, K. Michel, C. Mermel, S. J. Silver, B. A. Weir, J. H. Reiling, Q. Sheng, P. B. Gupta, R. C. Wadlow, H. Le, S. Hoersch, B. S. Wittner, S. Ramaswamy, D. M. Livingston, D. M. Sabatini, M. Meyerson, R. K. Thomas, E. S. Lander, J. P. Mesirov, D. E. Root, D. G. Gilliland, T. Jacks and W. C. Hahn (2009). "Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1." Nature **462**(7269): 108-112.

Barrero, M. J. and S. Malik (2006). "Two functional modes of a nuclear receptor-recruited arginine methyltransferase in transcriptional activation." Mol Cell **24**(2): 233-243.

Beattie, C. W., N. W. Hansen and P. A. Thomas (1985). "Steroid receptors in human lung cancer." Cancer Res **45**(9): 4206-4214.

Benesova, L., M. Minarik, D. Jancarikova, B. Belsanova and M. Pesek (2010). "Multiplicity of EGFR and KRAS mutations in non-small cell lung cancer

(NSCLC) patients treated with tyrosine kinase inhibitors." Anticancer Res **30**(5): 1667-1671.

Bernstein, E., A. A. Caudy, S. M. Hammond and G. J. Hannon (2001). "Role for a bidentate ribonuclease in the initiation step of RNA interference." Nature **409**(6818): 363-366.

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

Birmingham, A., E. M. Anderson, A. Reynolds, D. Ilsley-Tyree, D. Leake, Y. Fedorov, S. Baskerville, E. Maksimova, K. Robinson, J. Karpilow, W. S. Marshall and A. Khvorova (2006). "3' UTR seed matches, but not overall identity, are associated with RNAi off-targets." Nat Methods **3**(3): 199-204.

Boch, C., J. Kollmeier, A. Roth, S. Stephan-Falkenau, D. Misch, W. Gruning, T. T. Bauer and T. Mairinger (2013). "The frequency of EGFR and KRAS mutations in non-small cell lung cancer (NSCLC): routine screening data for central Europe from a cohort study." BMJ Open **3**(4).

Bookout, A. L., Y. Jeong, M. Downes, R. T. Yu, R. M. Evans and D. J. Mangelsdorf (2006). "Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network." Cell **126**(4): 789-799.

Bourzac, K. M., M. P. Rounseville, X. Zarate, V. S. Maddula, D. C. Henderson, J. A. Luckey, B. Seligmann and D. W. Galbraith (2011). "A high-density quantitative

nuclease protection microarray platform for high throughput analysis of gene expression." J Biotechnol **154**(1): 68-75.

Boutros, P. C., S. K. Lau, M. Pintilie, N. Liu, F. A. Shepherd, S. D. Der, M. S. Tsao, L. Z. Penn and I. Jurisica (2009). "Prognostic gene signatures for non-small-cell lung cancer." Proc Natl Acad Sci U S A **106**(8): 2824-2828.

Boyle, P. (1997). "Cancer, cigarette smoking and premature death in Europe: a review including the Recommendations of European Cancer Experts Consensus Meeting, Helsinki, October 1996." Lung Cancer **17**(1): 1-60.

Braasch, D. A., S. Jensen, Y. Liu, K. Kaur, K. Arar, M. A. White and D. R. Corey (2003). "RNA interference in mammalian cells by chemically-modified RNA." Biochemistry **42**(26): 7967-7975.

Brass, A. L., D. M. Dykxhoorn, Y. Benita, N. Yan, A. Engelman, R. J. Xavier, J. Lieberman and S. J. Elledge (2008). "Identification of host proteins required for HIV infection through a functional genomic screen." Science **319**(5865): 921-926.

Brodersen, P. and O. Voinnet (2009). "Revisiting the principles of microRNA target recognition and mode of action." Nat Rev Mol Cell Biol **10**(2): 141-148.

Brummelkamp, T. R., R. Bernards and R. Agami (2002). "A system for stable expression of short interfering RNAs in mammalian cells." Science **296**(5567): 550-553.

Bu, Y. and D. Cao (2012). "The origin of cancer stem cells." Front Biosci (Schol Ed) **4**: 819-830.

Budiman, M. E., J. L. Bubenik and D. M. Driscoll (2011). "Identification of a signature motif for the eIF4a3-SECIS interaction." Nucleic Acids Res **39**(17): 7730-7739.

Buehler, E., Y. C. Chen and S. Martin (2012). "C911: A bench-level control for sequence specific siRNA off-target effects." PLoS One **7**(12): e51942.

Buehler, E., A. A. Khan, S. Marine, M. Rajaram, A. Bahl, J. Burchard and M. Ferrer (2012). "siRNA off-target effects in genome-wide screens identify signaling pathway members." Sci Rep **2**: 428.

Burotto, M., A. Thomas, D. Subramaniam, G. Giaccone and A. Rajan (2014). "Biomarkers in Early-Stage Non-Small-Cell Lung Cancer: Current Concepts and Future Directions." J Thorac Oncol.

Bushman, F. D., N. Malani, J. Fernandes, I. D'Orso, G. Cagney, T. L. Diamond, H. Zhou, D. J. Hazuda, A. S. Espeseth, R. Konig, S. Bandyopadhyay, T. Ideker, S. P. Goff, N. J. Krogan, A. D. Frankel, J. A. Young and S. K. Chanda (2009). "Host cell factors in HIV replication: meta-analysis of genome-wide studies." PLoS Pathog **5**(5): e1000437.

Buttigliero, C., C. Monagheddu, P. Petroni, A. Saini, L. Dogliotti, G. Ciccone and A. Berruti (2011). "Prognostic role of vitamin d status and efficacy of vitamin D supplementation in cancer patients: a systematic review." Oncologist **16**(9): 1215-1227.

Cagle, P. T. and L. R. Chirieac (2012). "Advances in treatment of lung cancer with targeted therapy." Arch Pathol Lab Med **136**(5): 504-509.

Cairns, R. A., I. S. Harris and T. W. Mak (2011). "Regulation of cancer cell metabolism." Nat Rev Cancer **11**(2): 85-95.

Carmeliet, P. (2005). "Angiogenesis in life, disease and medicine." Nature **438**(7070): 932-936.

Carter, R. F. (2001). "BRCA1, BRCA2 and breast cancer: a concise clinical review." Clin Invest Med **24**(3): 147-157.

Chan, C. C., J. Dostie, M. D. Diem, W. Feng, M. Mann, J. Rappsilber and G. Dreyfuss (2004). "eIF4A3 is a novel component of the exon junction complex." RNA **10**(2): 200-209.

Chatterjee-Kishore, M. (2006). "From genome to phenome--RNAi library screening and hit characterization using signaling pathway analysis." Curr Opin Drug Discov Devel **9**(2): 231-239.

Chen, X., P. Liu and H. H. Chou (2013). "Whole-genome thermodynamic analysis reduces siRNA off-target effects." PLoS One **8**(3): e58326.

Chilvers, C. E. (1996). "Depot medroxyprogesterone acetate and breast cancer. A review of current knowledge." Drug Saf **15**(3): 212-218.

Chiu, Y. L. and T. M. Rana (2003). "siRNA function in RNAi: a chemical modification analysis." RNA **9**(9): 1034-1048.

Choung, S., Y. J. Kim, S. Kim, H. O. Park and Y. C. Choi (2006). "Chemical modification of siRNAs to improve serum stability without loss of efficacy." Biochem Biophys Res Commun **342**(3): 919-927.

Clemons, M., S. Danson and A. Howell (2002). "Tamoxifen ("Nolvadex"): a review." Cancer Treat Rev **28**(4): 165-180.

Collingwood, T. N., F. D. Urnov and A. P. Wolffe (1999). "Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription." J Mol Endocrinol **23**(3): 255-275.

Conzen, S. D. (2008). "Minireview: nuclear receptors and breast cancer." Mol Endocrinol **22**(10): 2215-2228.

Cooper, S. and S. G. Spiro (2006). "Small cell lung cancer: treatment review." Respirology **11**(3): 241-248.

Crosbie, E. J., M. H. Einstein, S. Franceschi and H. C. Kitchener (2013). "Human papillomavirus and cervical cancer." Lancet **382**(9895): 889-899.

Cullen, B. R. (2006). "Enhancing and confirming the specificity of RNAi experiments." Nat Methods **3**(9): 677-681.

Czauderna, F., M. Fechtner, S. Dames, H. Aygun, A. Klippel, G. J. Pronk, K. Giese and J. Kaufmann (2003). "Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells." Nucleic Acids Res **31**(11): 2705-2716.

Czech, B. and G. J. Hannon (2011). "Small RNA sorting: matchmaking for Argonautes." Nat Rev Genet **12**(1): 19-31.

da Cunha Santos, G., F. A. Shepherd and M. S. Tsao (2011). "EGFR mutations and lung cancer." Annu Rev Pathol **6**: 49-69.

Dahlgren, C., H. Y. Zhang, Q. Du, M. Grahn, G. Norstedt, C. Wahlestedt and Z. Liang (2008). "Analysis of siRNA specificity on targets with double-nucleotide mismatches." Nucleic Acids Res **36**(9): e53.

Dalhoff, K., J. Dancey, L. Astrup, T. Skovsgaard, K. J. Hamberg, F. J. Lofts, O. Rosmorduc, S. Erlinger, J. Bach Hansen, W. P. Steward, T. Skov, F. Burcharth and T. R. Evans (2003). "A phase II study of the vitamin D analogue Seocalcitol in patients with inoperable hepatocellular carcinoma." Br J Cancer **89**(2): 252-257.

Das, S., S. Ghosal, K. Kozak and J. Chakrabarti (2013). "An siRNA designing tool with a unique functional off-target filtering approach." J Biomol Struct Dyn **31**(11): 1343-1357.

David, A. R. and M. R. Zimmerman (2010). "Cancer: an old disease, a new disease or something in between?" Nat Rev Cancer **10**(10): 728-733.

Dedhia, P., S. Tarale, G. Dhongde, R. Khadapkar and B. Das (2007). "Evaluation of DNA extraction methods and real time PCR optimization on formalin-fixed paraffin-embedded tissues." Asian Pac J Cancer Prev **8**(1): 55-59.

Delyani, J. A. (2000). "Mineralocorticoid receptor antagonists: the evolution of utility and pharmacology." Kidney Int **57**(4): 1408-1411.

Ding, L., M. Paszkowski-Rogacz, A. Nitzsche, M. M. Slabicki, A. K. Heninger, I. de Vries, R. Kittler, M. Junqueira, A. Shevchenko, H. Schulz, N. Hubner, M. X. Doss, A. Sachinidis, J. Hescheler, R. Iacone, K. Anastassiadis, A. F. Stewart, M.

T. Pisabarro, A. Caldarelli, I. Poser, M. Theis and F. Buchholz (2009). "A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity." Cell Stem Cell **4**(5): 403-415.

Ding, L. H., Y. Xie, S. Park, G. Xiao and M. D. Story (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**(10): e58.

Doench, J. G., C. P. Petersen and P. A. Sharp (2003). "siRNAs can function as miRNAs." Genes Dev **17**(4): 438-442.

Dragnev, K., M. You, Y. Wang and R. Lubet (2013). "Lung cancer chemoprevention: difficulties, promise and potential agents?" Expert Opin Investig Drugs **22**(1): 35-47.

Du, Q., H. Thonberg, J. Wang, C. Wahlestedt and Z. Liang (2005). "A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites." Nucleic Acids Res **33**(5): 1671-1677.

Echeverri, C. J., P. A. Beachy, B. Baum, M. Boutros, F. Buchholz, S. K. Chanda, J. Downward, J. Ellenberg, A. G. Fraser, N. Hacohen, W. C. Hahn, A. L. Jackson, A. Kiger, P. S. Linsley, L. Lum, Y. Ma, B. Mathey-Prevot, D. E. Root, D. M. Sabatini, J. Taipale, N. Perrimon and R. Bernards (2006). "Minimizing the risk of reporting false positives in large-scale RNAi screens." Nat Methods **3**(10): 777-779.

Echeverri, C. J. and N. Perrimon (2006). "High-throughput RNAi screening in cultured cells: a user's guide." Nat Rev Genet **7**(5): 373-384.

Eden, E., R. Navon, I. Steinfeld, D. Lipson and Z. Yakhini (2009). "GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists." BMC Bioinformatics **10**: 48.

Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl (2001). "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells." Nature **411**(6836): 494-498.

Elbashir, S. M., J. Martinez, A. Patkaniowska, W. Lendeckel and T. Tuschl (2001). "Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate." EMBO J **20**(23): 6877-6888.

Elmen, J., H. Thonberg, K. Ljungberg, M. Frieden, M. Westergaard, Y. Xu, B. Wahren, Z. Liang, H. Orum, T. Koch and C. Wahlestedt (2005). "Locked nucleic acid (LNA) mediated improvements in siRNA stability and functionality." Nucleic Acids Res **33**(1): 439-447.

Engels, J. W. (2013). "Gene silencing by chemically modified siRNAs." N Biotechnol **30**(3): 302-307.

Engstrom, M. J., S. Opdahl, A. I. Hagen, P. R. Romundstad, L. A. Akslen, O. A. Haugen, L. J. Vatten and A. M. Bofin (2013). "Molecular subtypes, histopathological grade and survival in a historic cohort of breast cancer patients." Breast Cancer Res Treat **140**(3): 463-473.

Erickson, H. S. (2012). "Measuring molecular biomarkers in epidemiologic studies: laboratory techniques and biospecimen considerations." Stat Med **31**(22): 2400-2413.

Escriva, H., M. C. Langlois, R. L. Mendonca, R. Pierce and V. Laudet (1998). "Evolution and diversification of the nuclear receptor superfamily." Ann N Y Acad Sci **839**: 143-146.

Evans, T. R., K. W. Colston, F. J. Lofts, D. Cunningham, D. A. Anthoney, H. Gogas, J. S. de Bono, K. J. Hamberg, T. Skov and J. L. Mansi (2002). "A phase II trial of the vitamin D analogue Seocalcitol (EB1089) in patients with inoperable pancreatic cancer." Br J Cancer **86**(5): 680-685.

Farragher, S. M., A. Tanney, R. D. Kennedy and D. Paul Harkin (2008). "RNA expression analysis from formalin fixed paraffin embedded tissues." Histochem Cell Biol **130**(3): 435-445.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello (1998). "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans." Nature **391**(6669): 806-811.

Fritz, H., D. Kennedy, D. Fergusson, R. Fernandes, S. Doucette, K. Cooley, A. Seely, S. Sagar, R. Wong and D. Seely (2011). "Vitamin A and retinoid derivatives for lung cancer: a systematic review and meta analysis." PLoS One **6**(6): e21107.

Gambertoglio, J. G., W. J. Amend, Jr. and L. Z. Benet (1980). "Pharmacokinetics and bioavailability of prednisone and prednisolone in healthy volunteers and patients: a review." J Pharmacokinet Biopharm **8**(1): 1-52.

Gehring, N. H., S. Lamprinaki, M. W. Hentze and A. E. Kulozik (2009). "The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay." PLoS Biol **7**(5): e1000120.

Germain, P., J. Iyer, C. Zechel and H. Gronemeyer (2002). "Co-regulator recruitment and the mechanism of retinoic acid receptor synergy." Nature **415**(6868): 187-192.

Germain, P., B. Staels, C. Dacquet, M. Spedding and V. Laudet (2006). "Overview of nomenclature of nuclear receptors." Pharmacol Rev **58**(4): 685-704.

Gerson, K. D., V. S. Maddula, B. E. Seligmann, J. R. Shearstone, A. Khan and A. M. Mercurio (2012). "Effects of beta4 integrin expression on microRNA patterns in breast cancer." Biol Open **1**(7): 658-666.

Gillatt, D. (2006). "Antiandrogen treatments in locally advanced prostate cancer: are they all the same?" J Cancer Res Clin Oncol **132 Suppl 1**: S17-26.

Glass, C. K. and M. G. Rosenfeld (2000). "The coregulator exchange in transcriptional functions of nuclear receptors." Genes Dev **14**(2): 121-141.

Gloeckler Ries, L. A., M. E. Reichman, D. R. Lewis, B. F. Hankey and B. K. Edwards (2003). "Cancer survival and incidence from the Surveillance, Epidemiology, and End Results (SEER) program." Oncologist **8**(6): 541-552.

Gnanapragasam, V. J. (2010). "Unlocking the molecular archive: the emerging use of formalin-fixed paraffin-embedded tissue for biomarker research in urological cancer." BJU Int **105**(2): 274-278.

Goffin, J., C. Lacchetti, P. M. Ellis, Y. C. Ung and W. K. Evans (2010). "First-line systemic chemotherapy in the treatment of advanced non-small cell lung cancer: a systematic review." J Thorac Oncol **5**(2): 260-274.

Gonzalez-Angulo, A. M., J. Ferrer-Lozano, K. Stemke-Hale, A. Sahin, S. Liu, J. A. Barrera, O. Burgues, A. M. Lluch, H. Chen, G. N. Hortobagyi, G. B. Mills and F. Meric-Bernstam (2011). "PI3K pathway mutations and PTEN levels in primary and metastatic breast cancer." Mol Cancer Ther **10**(6): 1093-1101.

Greenberg, A. K., J. Hu, S. Basu, J. Hay, J. Reibman, T. A. Yie, K. M. Tchou-Wong, W. N. Rom and T. C. Lee (2002). "Glucocorticoids inhibit lung cancer cell growth through both the extracellular signal-related kinase pathway and cell cycle regulators." Am J Respir Cell Mol Biol **27**(3): 320-328.

Greschik, H., R. Flaig, J. P. Renaud and D. Moras (2004). "Structural basis for the deactivation of the estrogen-related receptor gamma by diethylstilbestrol or 4-hydroxytamoxifen and determinants of selectivity." J Biol Chem **279**(32): 33639-33646.

Greystoke, A. and S. A. Mullamitha (2012). "How many diseases are colorectal cancer?" Gastroenterol Res Pract **2012**: 564741.

Grimson, A., K. K. Farh, W. K. Johnston, P. Garrett-Engele, L. P. Lim and D. P. Bartel (2007). "MicroRNA targeting specificity in mammals: determinants beyond seed pairing." Mol Cell **27**(1): 91-105.

Grulich, A. E., M. T. van Leeuwen, M. O. Falster and C. M. Vajdic (2007). "Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis." Lancet **370**(9581): 59-67.

Grundker, C., A. R. Gunthert and G. Emons (2008). "Hormonal heterogeneity of endometrial cancer." Adv Exp Med Biol **630**: 166-188.

Gu, S., L. Jin, Y. Zhang, Y. Huang, F. Zhang, P. N. Valdmanis and M. A. Kay (2012). "The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo." Cell **151**(4): 900-911.

Gyorffy, B., P. Surowiak, J. Budczies and A. Lanczky (2013). "Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer." PLoS One **8**(12): e82241.

Hajdu, S. I. (2011). "A note from history: landmarks in history of cancer, part 1." Cancer **117**(5): 1097-1102.

Hamilton, A. J. and D. C. Baulcombe (1999). "A species of small antisense RNA in posttranscriptional gene silencing in plants." Science **286**(5441): 950-952.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.

Hantzis, V., A. Albiston, D. Matsacos, E. M. Wintour, A. Peers, I. Koukoulas, K. Myles, K. Moritz and M. Dodic (2002). "Effect of early glucocorticoid treatment on MR and GR in late gestation ovine kidney." Kidney Int **61**(2): 405-413.

Harousseau, J. L., M. Attal, X. Leleu, J. Troncy, B. Pegourie, A. M. Stoppa, C. Hulin, L. Benboubker, J. G. Fuzibet, M. Renaud, P. Moreau and H. Avet-Loiseau (2006). "Bortezomib plus dexamethasone as induction treatment prior to autologous stem cell transplantation in patients with newly diagnosed multiple myeloma: results of an IFM phase II study." Haematologica **91**(11): 1498-1505.

Heinlein, C. A. and C. Chang (2004). "Androgen receptor in prostate cancer." Endocr Rev **25**(2): 276-308.

Henschke, C. I., D. F. Yankelevitz, D. M. Libby, M. W. Pasmantier, J. P. Smith and O. S. Miettinen (2006). "Survival of patients with stage I lung cancer detected on CT screening." N Engl J Med **355**(17): 1763-1771.

Hollstein, M., D. Sidransky, B. Vogelstein and C. C. Harris (1991). "p53 mutations in human cancers." Science **253**(5015): 49-53.

Huang, P., V. Chandra and F. Rastinejad (2010). "Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics." Annu Rev Physiol **72**: 247-272.

Inoue, K., Y. Kawahito, Y. Tsubouchi, R. Yamada, M. Kohno, Y. Hosokawa, D. Katoh, D. Bishop-Bailey, T. Hla and H. Sano (2001). "Expression of peroxisome

proliferator-activated receptor (PPAR)-gamma in human lung cancer." Anticancer Res **21**(4A): 2471-2476.

Iyer, S., K. Deutsch, X. Yan and B. Lin (2007). "Batch RNAi selector: a standalone program to predict specific siRNA candidates in batches with enhanced sensitivity." Comput Methods Programs Biomed **85**(3): 203-209.

Jackson, A. L., S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley (2003). "Expression profiling reveals off-target gene regulation by RNAi." Nat Biotechnol **21**(6): 635-637.

Jackson, A. L., J. Burchard, D. Leake, A. Reynolds, J. Schelter, J. Guo, J. M. Johnson, L. Lim, J. Karpilow, K. Nichols, W. Marshall, A. Khvorova and P. S. Linsley (2006). "Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing." RNA **12**(7): 1197-1205.

Jackson, A. L., J. Burchard, J. Schelter, B. N. Chau, M. Cleary, L. Lim and P. S. Linsley (2006). "Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity." RNA **12**(7): 1179-1187.

Jackson, A. L. and L. A. Loeb (2001). "The contribution of endogenous sources of DNA damage to the multiple mutations in cancer." Mutat Res **477**(1-2): 7-21.

Jeal, W. and D. Faulds (1997). "Triamcinolone acetonide. A review of its pharmacological properties and therapeutic efficacy in the management of allergic rhinitis." Drugs **53**(2): 257-280.

Jemal, A., F. Bray, M. M. Center, J. Ferlay, E. Ward and D. Forman (2011). "Global cancer statistics." CA Cancer J Clin **61**(2): 69-90.

Jeong, Y., Y. Xie, W. Lee, A. L. Bookout, L. Girard, G. Raso, C. Behrens, Wistuba, II, A. F. Gadzar, J. D. Minna and D. J. Mangelsdorf (2012). "Research resource: Diagnostic and therapeutic potential of nuclear receptor expression in lung cancer." Mol Endocrinol **26**(8): 1443-1454.

Jeong, Y., Y. Xie, G. Xiao, C. Behrens, L. Girard, Wistuba, II, J. D. Minna and D. J. Mangelsdorf (2010). "Nuclear receptor expression defines a set of prognostic biomarkers for lung cancer." PLoS Med **7**(12): e1000378.

Jeong, Y., Y. Xie, G. Xiao, C. Behrens, L. Girard, I. I. Wistuba, J. D. Minna and D. J. Mangelsdorf (2010). "Nuclear Receptor Expression Defines a Set of Prognostic Biomarkers for Lung Cancer." PLoS Med **7**(12): e1000378.

Jones, S., X. Zhang, D. W. Parsons, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, S. M. Hong, B. Fu, M. T. Lin, E. S. Calhoun, M. Kamiyama, K. Walter, T. Nikolskaya, Y. Nikolsky, J. Hartigan, D. R. Smith, M. Hidalgo, S. D. Leach, A. P. Klein, E. M. Jaffee, M. Goggins, A. Maitra, C. Iacobuzio-Donahue, J. R. Eshleman, S. E. Kern, R. H. Hruban, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu and K. W. Kinzler (2008). "Core signaling pathways in human pancreatic cancers revealed by global genomic analyses." Science **321**(5897): 1801-1806.

Kaiser, U., J. Hofmann, M. Schilli, B. Wegmann, U. Klotz, S. Wedel, A. K. Virmani, E. Wollmer, D. Branscheid, A. F. Gazdar and K. Havemann (1996). "Steroid-hormone receptors in cell lines and tumor biopsies of human lung cancer." Int J Cancer **67**(3): 357-364.

Kamangar, F., G. M. Dores and W. F. Anderson (2006). "Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world." J Clin Oncol **24**(14): 2137-2150.

Kashima, I., A. Yamashita, N. Izumi, N. Kataoka, R. Morishita, S. Hoshino, M. Ohno, G. Dreyfuss and S. Ohno (2006). "Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay." Genes Dev **20**(3): 355-367.

Katkoori, V. R., C. Shanmugam, X. Jia, S. P. Vitta, M. Sthanam, T. Callens, L. Messiaen, D. Chen, B. Zhang, H. L. Bumpers, T. Samuel and U. Manne (2012). "Prognostic significance and gene expression profiles of p53 mutations in microsatellite-stable stage III colorectal adenocarcinomas." PLoS One **7**(1): e30020.

Kaur, H., A. Arora, J. Wengel and S. Maiti (2006). "Thermodynamic, counterion, and hydration effects for the incorporation of locked nucleic acid nucleotides into DNA duplexes." Biochemistry **45**(23): 7347-7355.

Kay, P., G. Schlossmacher, L. Matthews, P. Sommer, D. Singh, A. White and D. Ray (2011). "Loss of glucocorticoid receptor expression by DNA methylation prevents

glucocorticoid induced apoptosis in human small cell lung cancer cells." PLoS One **6**(10): e24839.

Keshamouni, V. G., S. Han and J. Roman (2007). "Peroxisome proliferator-activated receptors in lung cancer." PPAR Res **2007**: 90289.

Khuri, F. R., R. Lotan, B. L. Kemp, S. M. Lippman, H. Wu, L. Feng, J. J. Lee, C. S. Cooksley, B. Parr, E. Chang, G. L. Walsh, J. S. Lee, W. K. Hong and X. C. Xu (2000). "Retinoic acid receptor-beta as a prognostic indicator in stage I non-small-cell lung cancer." J Clin Oncol **18**(15): 2798-2804.

Khuri, F. R., J. R. Rigas, R. A. Figlin, R. J. Gralla, D. M. Shin, R. Munden, N. Fox, M. R. Huyghe, Y. Kean, S. D. Reich and W. K. Hong (2001). "Multi-institutional phase I/II trial of oral bexarotene in combination with cisplatin and vinorelbine in previously untreated patients with advanced non-small-cell lung cancer." J Clin Oncol **19**(10): 2626-2637.

Kim, H. S., S. Mendiratta, J. Kim, C. V. Pecot, J. E. Larsen, I. Zubovych, B. Y. Seo, B. Eskiocak, H. Chung, E. McMillan, S. Wu, J. De Brabander, K. Komurov, J. E. Toombs, S. Wei, M. Peyton, N. Williams, A. F. Gazdar, B. A. Posner, R. A. Brekken, A. K. Sood, R. J. Deberardinis, M. G. Roth, J. D. Minna and M. A. White (2013). "Systematic identification of molecular subtype-selective vulnerabilities in non-small-cell lung cancer." Cell **155**(3): 552-566.

Kim, S. H., G. Chen, A. N. King, C. K. Jeon, P. J. Christensen, L. Zhao, R. U. Simpson, D. G. Thomas, T. J. Giordano, D. E. Brenner, B. Hollis, D. G. Beer and N.

Ramnath (2012). "Characterization of vitamin D receptor (VDR) in lung adenocarcinoma." Lung Cancer **77**(2): 265-271.

Kim, Y., M. Koh, D. K. Kim, H. S. Choi and S. B. Park (2009). "Efficient discovery of selective small molecule agonists of estrogen-related receptor gamma using combinatorial approach." J Comb Chem **11**(5): 928-937.

King, H. A., K. W. Trotter and T. K. Archer (2012). "Chromatin remodeling during glucocorticoid receptor regulated transactivation." Biochim Biophys Acta **1819**(7): 716-726.

Kittler, R., L. Pelletier, A. K. Heninger, M. Slabicki, M. Theis, L. Miroslaw, I. Poser, S. Lawo, H. Grabner, K. Kozak, J. Wagner, V. Surendranath, C. Richter, W. Bowen, A. L. Jackson, B. Habermann, A. A. Hyman and F. Buchholz (2007). "Genome-scale RNAi profiling of cell division in human tissue culture cells." Nat Cell Biol **9**(12): 1401-1412.

Kittler, R., V. Surendranath, A. K. Heninger, M. Slabicki, M. Theis, G. Putz, K. Franke, A. Caldarelli, H. Grabner, K. Kozak, J. Wagner, E. Rees, B. Korn, C. Frenzel, C. Sachse, B. Sonnichsen, J. Guo, J. Schelter, J. Burchard, P. S. Linsley, A. L. Jackson, B. Habermann and F. Buchholz (2007). "Genome-wide resources of endoribonuclease-prepared short interfering RNAs for specific loss-of-function studies." Nat Methods **4**(4): 337-344.

Kittler, R., J. Zhou, S. Hua, L. Ma, Y. Liu, E. Pendleton, C. Cheng, M. Gerstein and K. P. White (2013). "A comprehensive nuclear receptor network for breast cancer cells." Cell Rep **3**(2): 538-551.

Klinge, C. M. (2012). "Inhibition of non-small-cell lung cancer growth by combined fulvestrant and vandetanib." Future Oncol **8**(5): 529-533.

Klinghoffer, R. A., J. Magnus, J. Schelter, M. Mehaffey, C. Coleman and M. A. Cleary (2010). "Reduced seed region-based off-target activity with lentivirus-mediated RNAi." RNA **16**(5): 879-884.

Knudson, A. G. (2001). "Two genetic hits (more or less) to cancer." Nat Rev Cancer **1**(2): 157-162.

Kolas, N. K., J. R. Chapman, S. Nakada, J. Ylanko, R. Chahwan, F. D. Sweeney, S. Panier, M. Mendez, J. Wildenhain, T. M. Thomson, L. Pelletier, S. P. Jackson and D. Durocher (2007). "Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase." Science **318**(5856): 1637-1640.

Kraus, W. L. and J. Wong (2002). "Nuclear receptor-dependent transcription with chromatin. Is it all about enzymes?" Eur J Biochem **269**(9): 2275-2283.

Kraynack, B. A. and B. F. Baker (2006). "Small interfering RNAs containing full 2'-O-methylribonucleotide-modified sense strands display Argonaute2/eIF2C2-dependent activity." RNA **12**(1): 163-176.

Kris, R. M., S. Felder, M. Deyholos, G. M. Lambert, J. Hinton, I. Botros, R. Martel, B. Seligmann and D. W. Galbraith (2007). "High-throughput, high-sensitivity analysis of gene expression in Arabidopsis." Plant Physiol **144**(3): 1256-1266.

Kumar, R. and E. B. Thompson (1999). "The structure of the nuclear hormone receptors." Steroids **64**(5): 310-319.

Larsen, J. E., T. Cascone, D. E. Gerber, J. V. Heymach and J. D. Minna (2011). "Targeted therapies for lung cancer: clinical experience and novel agents." Cancer J **17**(6): 512-527.

Larsen, J. E., S. J. Pavey, L. H. Passmore, R. Bowman, B. E. Clarke, N. K. Hayward and K. M. Fong (2007). "Expression profiling defines a recurrence signature in lung squamous cell carcinoma." Carcinogenesis **28**(3): 760-766.

Larson, M. H., L. A. Gilbert, X. Wang, W. A. Lim, J. S. Weissman and L. S. Qi (2013). "CRISPR interference (CRISPRi) for sequence-specific control of gene expression." Nat. Protocols **8**(11): 2180-2196.

Lawo, S., M. Bashkurov, M. Mullin, M. G. Ferreria, R. Kittler, B. Habermann, A. Tagliaferro, I. Poser, J. R. Hutchins, B. Hegemann, D. Pinchev, F. Buchholz, J. M. Peters, A. A. Hyman, A. C. Gingras and L. Pelletier (2009). "HAUS, the 8-subunit human Augmin complex, regulates centrosome and spindle integrity." Curr Biol **19**(10): 816-826.

Leber, M. F. and T. Efferth (2009). "Molecular principles of cancer invasion and metastasis (review)." Int J Oncol **34**(4): 881-895.

Lee, J. W., Y. C. Lee, S. Y. Na, D. J. Jung and S. K. Lee (2001). "Transcriptional coregulators of the nuclear receptor superfamily: coactivators and corepressors." Cell Mol Life Sci **58**(2): 289-297.

Lee, Y. H., S. S. Koh, X. Zhang, X. Cheng and M. R. Stallcup (2002). "Synergy among nuclear receptor coactivators: selective requirement for protein methyltransferase and acetyltransferase activities." Mol Cell Biol **22**(11): 3621-3632.

Lengauer, C., K. W. Kinzler and B. Vogelstein (1998). "Genetic instabilities in human cancers." Nature **396**(6712): 643-649.

Lewis, B. P., C. B. Burge and D. P. Bartel (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.

Li, M. Y., A. W. Kong, H. Yuan, L. T. Ma, M. K. Hsin, I. Y. Wan, M. J. Underwood and G. G. Chen (2012). "Pioglitazone prevents smoking carcinogen-induced lung tumor development in mice." Curr Cancer Drug Targets **12**(6): 597-606.

Li, M. Y., T. W. Lee, A. P. Yim and G. G. Chen (2006). "Function of PPARgamma and its ligands in lung cancer." Crit Rev Clin Lab Sci **43**(2): 183-202.

Liang, H., P. Kowalczyk, J. J. Junco, H. L. Klug-De Santiago, G. Malik, S. J. Wei and T. J. Slaga (2014). "Differential effects on lung cancer cell proliferation by agonists of glucocorticoid and PPARalpha receptors." Mol Carcinog **53**(9): 753-763.

Lin, X., X. Ruan, M. G. Anderson, J. A. McDowell, P. E. Kroeger, S. W. Fesik and Y. Shen (2005). "siRNA-mediated off-target gene silencing triggered by a 7 nt complementation." Nucleic Acids Res **33**(14): 4527-4535.

Liu, W. Y., Y. Wang, Y. H. Sun, Y. P. Wang, S. P. Chen and Z. Y. Zhu (2005). "Efficient RNA interference in zebrafish embryos using siRNA synthesized with SP6 RNA polymerase." Dev Growth Differ **47**(5): 323-331.

Loeb, L. A., V. L. Ernster, K. E. Warner, J. Abbotts and J. Laszlo (1984). "Smoking and lung cancer: an overview." Cancer Res **44**(12 Pt 1): 5940-5958.

Lonard, D. M., Z. Nawaz, C. L. Smith and B. W. O'Malley (2000). "The 26S proteasome is required for estrogen receptor-alpha and coactivator turnover and for efficient estrogen receptor-alpha transactivation." Mol Cell **5**(6): 939-948.

Luo, J., M. J. Emanuele, D. Li, C. J. Creighton, M. R. Schlabach, T. F. Westbrook, K. K. Wong and S. J. Elledge (2009). "A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene." Cell **137**(5): 835-848.

Ma, Y., A. Creanga, L. Lum and P. A. Beachy (2006). "Prevalence of off-target effects in Drosophila RNA interference screens." Nature **443**(7109): 359-363.

Manavathi, B. and R. Kumar (2007). "Metastasis tumor antigens, an emerging family of multifaceted master coregulators." J Biol Chem **282**(3): 1529-1533.

Mangelsdorf, D. J., C. Thummel, M. Beato, P. Herrlich, G. Schutz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon and R. M. Evans (1995). "The nuclear receptor superfamily: the second decade." Cell **83**(6): 835-839.

Martel, R. R., I. W. Botros, M. P. Rounseville, J. P. Hinton, R. R. Staples, D. A. Morales, J. B. Farmer and B. E. Seligmann (2002). "Multiplexed screening assay for mRNA combining nuclease protection with luminescent array detection." Assay Drug Dev Technol **1**(1 Pt 1): 61-71.

McKenna, N. J., R. B. Lanz and B. W. O'Malley (1999). "Nuclear receptor coregulators: cellular and molecular biology." Endocr Rev **20**(3): 321-344.

McKenna, N. J., Z. Nawaz, S. Y. Tsai, M. J. Tsai and B. W. O'Malley (1998). "Distinct steady-state nuclear receptor coregulator complexes exist in vivo." Proc Natl Acad Sci U S A **95**(20): 11697-11702.

McKenna, N. J. and B. W. O'Malley (2002). "Combinatorial control of gene expression by nuclear receptors and coregulators." Cell **108**(4): 465-474.

Mikkonen, L., P. Pihlajamaa, B. Sahu, F. P. Zhang and O. A. Janne (2010). "Androgen receptor and androgen-dependent gene expression in lung." Mol Cell Endocrinol **317**(1-2): 14-24.

Miner, J. N., M. H. Hong and A. Negro-Vilar (2005). "New and improved glucocorticoid receptor ligands." Expert Opin Investig Drugs **14**(12): 1527-1545.

Mohr, S., C. Bakal and N. Perrimon (2010). "Genomic screening with RNAi: results and challenges." Annu Rev Biochem **79**: 37-64.

Mook, O. R., F. Baas, M. B. de Wissel and K. Fluiter (2007). "Evaluation of locked nucleic acid-modified small interfering RNA in vitro and in vivo." Mol Cancer Ther **6**(3): 833-843.

Murray, A. W. (2004). "Recycling the cell cycle: cyclins revisited." Cell **116**(2): 221-234.

Myers, J. W., J. T. Chi, D. Gong, M. E. Schaner, P. O. Brown and J. E. Ferrell (2006). "Minimizing off-target effects by using diced siRNAs for RNA interference." J RNAi Gene Silencing **2**(2): 181-194.

Mysara, M., J. M. Garibaldi and M. Elhefnawi (2011). "MysiRNA-designer: a workflow for efficient siRNA design." PLoS One **6**(10): e25642.

Naito, Y. and K. Ui-Tei (2012). "siRNA Design Software for a Target Gene-Specific RNA Interference." Front Genet **3**: 102.

Nemenoff, R. A. and R. A. Winn (2005). "Role of nuclear receptors in lung tumourigenesis." Eur J Cancer **41**(16): 2561-2568.

Newman, J. T., M. D. Nellermoe and J. L. Carnett (1992). "Hydrocortisone phonophoresis. A literature review." J Am Podiatr Med Assoc **82**(8): 432-435.

Ngoan, L. T., T. Mizoue, Y. Fujino, N. Tokui and T. Yoshimura (2002). "Dietary factors and stomach cancer mortality." Br J Cancer **87**(1): 37-42.

Nojima, H. (1997). "Cell cycle checkpoints, chromosome stability and the progression of cancer." Hum Cell **10**(4): 221-230.

O'Malley, B. W. (2006). "Molecular biology. Little molecules with big goals." Science **313**(5794): 1749-1750.

O'Malley, B. W. and N. J. McKenna (2008). "Coactivators and corepressors: what's in a name?" Mol Endocrinol **22**(10): 2213-2214.

O'Malley, B. W. and R. Kumar (2009). "Nuclear Receptor Coregulators in Cancer Biology." Cancer research **69**(21): 8217-8222.

Olejniczak, M., P. Galka and W. J. Krzyzosiak (2010). "Sequence-non-specific effects of RNA interference triggers and microRNA regulators." Nucleic Acids Res **38**(1): 1-16.

Otrock, Z. K., R. A. Mahfouz, J. A. Makarem and A. I. Shamseddine (2007). "Understanding the biology of angiogenesis: review of the most important molecular mechanisms." Blood Cells Mol Dis **39**(2): 212-220.

Overington, J. P., B. Al-Lazikani and A. L. Hopkins (2006). "How many drug targets are there?" Nat Rev Drug Discov **5**(12): 993-996.

Owczarzy, R., Y. You, C. L. Groth and A. V. Tataurov (2011). "Stability and mismatch discrimination of locked nucleic acid-DNA duplexes." Biochemistry **50**(43): 9352-9367.

Paddison, P. J., A. A. Caudy, E. Bernstein, G. J. Hannon and D. S. Conklin (2002). "Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells." Genes Dev **16**(8): 948-958.

Palles, C., J. B. Cazier, K. M. Howarth, E. Domingo, A. M. Jones, P. Broderick, Z. Kemp, S. L. Spain, E. Guarino, I. Salguero, A. Sherborne, D. Chubb, L. G. Carvajal-Carmona, Y. Ma, K. Kaur, S. Dobbins, E. Barclay, M. Gorman, L. Martin, M. B. Kovac, S. Humphray, A. Lucassen, C. C. Holmes, D. Bentley, P. Donnelly, J. Taylor, C. Petridis, R. Roylance, E. J. Sawyer, D. J. Kerr, S. Clark, J.

Grimes, S. E. Kearsey, H. J. Thomas, G. McVean, R. S. Houlston and I. Tomlinson (2013). "Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas." Nat Genet **45**(2): 136-144.

Pao, W. and N. Girard (2011). "New driver mutations in non-small-cell lung cancer." Lancet Oncol **12**(2): 175-180.

Parsons, B. D., A. Schindler, D. H. Evans and E. Foley (2009). "A direct phenotypic comparison of siRNA pools and multiple individual duplexes in a functional assay." PLoS One **4**(12): e8471.

Pechhold, S., M. Stouffer, G. Walker, R. Martel, B. Seligmann, Y. Hang, R. Stein, D. M. Harlan and K. Pechhold (2009). "Transcriptional analysis of intracytoplasmically stained, FACS-purified cells by high-throughput, quantitative nuclease protection." Nat Biotechnol **27**(11): 1038-1042.

Perez, E. A., D. R. Gandara, M. J. Edelman, R. O'Donnell, I. J. Lauder and M. DeGregorio (2003). "Phase I trial of high-dose tamoxifen in combination with cisplatin in patients with lung cancer and other advanced malignancies." Cancer Invest **21**(1): 1-6.

Peterson, C., T. J. Hursti, S. Borjeson, E. Avall-Lundqvist, M. Fredrikson, C. J. Furst, H. Lomberg and G. Steineck (1996). "Single high-dose dexamethasone improves the effect of ondansetron on acute chemotherapy-induced nausea and vomiting but impairs the control of delayed symptoms." Support Care Cancer **4**(6): 440-446.

Phelan, M. C. (2007). "Basic techniques in mammalian cell tissue culture." Curr Protoc Cell Biol **Chapter 1**: Unit 1 1.

Philip, M., D. A. Rowley and H. Schreiber (2004). "Inflammation as a tumor promoter in cancer induction." Semin Cancer Biol **14**(6): 433-439.

Rabbani, F. and C. Cordon-Cardo (2000). "Mutation of cell cycle regulators and their impact on superficial bladder cancer." Urol Clin North Am **27**(1): 83-102, ix.

Ramnath, N., S. Daignault-Newton, G. K. Dy, J. R. Muindi, A. Adjei, V. L. Elingrod, G. P. Kalemkerian, K. B. Cease, P. J. Stella, D. E. Brenner, S. Troeschel, C. S. Johnson and D. L. Trump (2013). "A phase I/II pharmacokinetic and pharmacogenomic study of calcitriol in combination with cisplatin and docetaxel in advanced non-small-cell lung cancer." Cancer Chemother Pharmacol **71**(5): 1173-1182.

Rand, T. A., S. Petersen, F. Du and X. Wang (2005). "Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation." Cell **123**(4): 621-629.

Rao, D. D., N. Senzer, M. A. Cleary and J. Nemunaitis (2009). "Comparative assessment of siRNA and shRNA off target effects: what is slowing clinical development." Cancer Gene Ther **16**(11): 807-809.

Rao, D. D., J. S. Vorhies, N. Senzer and J. Nemunaitis (2009). "siRNA vs. shRNA: similarities and differences." Adv Drug Deliv Rev **61**(9): 746-759.

Recchia, A. G., A. M. Musti, M. Lanzino, M. L. Panno, E. Turano, R. Zumpano, A. Belfiore, S. Ando and M. Maggiolini (2009). "A cross-talk between the androgen

receptor and the epidermal growth factor receptor leads to p38MAPK-dependent activation of mTOR and cyclinD1 expression in prostate and lung cancer cells." Int J Biochem Cell Biol **41**(3): 603-614.

Richard-Foy, H., F. D. Sistare, A. T. Riegel, S. S. Simons, Jr. and G. L. Hager (1987). "Mechanism of dexamethasone 21-mesylate antiglucocorticoid action: II. Receptor-antiglucocorticoid complexes do not interact productively with mouse mammary tumor virus long terminal repeat chromatin." Mol Endocrinol **1**(9): 659-665.

Rimsza, L. M., M. L. Leblanc, J. M. Unger, T. P. Miller, T. M. Grogan, D. O. Persky, R. R. Martel, C. M. Sabalos, B. Seligmann, R. M. Braziel, E. Campo, A. Rosenwald, J. M. Connors, L. H. Sehn, N. Johnson and R. D. Gascoyne (2008). "Gene expression predicts overall survival in paraffin-embedded tissues of diffuse large B-cell lymphoma treated with R-CHOP." Blood **112**(8): 3425-3433.

Rimsza, L. M., G. Wright, M. Schwartz, W. C. Chan, E. S. Jaffe, R. D. Gascoyne, E. Campo, A. Rosenwald, G. Ott, J. R. Cook, R. R. Tubbs, R. M. Braziel, J. Delabie, T. P. Miller and L. M. Staudt (2011). "Accurate classification of diffuse large B-cell lymphoma into germinal center and activated B-cell subtypes using a nuclease protection assay on formalin-fixed, paraffin-embedded tissues." Clin Cancer Res **17**(11): 3727-3732.

Roberts, P. J. and T. E. Stinchcombe (2013). "KRAS mutation: should we test for it, and does it matter?" J Clin Oncol **31**(8): 1112-1121.

Roberts, R. A., C. M. Sabalos, M. L. LeBlanc, R. R. Martel, Y. M. Frutiger, J. M. Unger, I. W. Botros, M. P. Rounseville, B. E. Seligmann, T. P. Miller, T. M. Grogan and L. M. Rimsza (2007). "Quantitative nuclease protection assay in paraffin-embedded tissue replicates prognostic microarray gene expression in diffuse large-B-cell lymphoma." Lab Invest **87**(10): 979-997.

Rowan, B. G., N. L. Weigel and B. W. O'Malley (2000). "Phosphorylation of steroid receptor coactivator-1. Identification of the phosphorylation sites and phosphorylation through the mitogen-activated protein kinase pathway." J Biol Chem **275**(6): 4475-4483.

Safe, S., U. H. Jin, E. Hedrick, A. Reeder and S. O. Lee (2014). "Minireview: role of orphan nuclear receptors in cancer and potential as drug targets." Mol Endocrinol **28**(2): 157-172.

Sanchez-Cespedes, M. (2011). "The role of LKB1 in lung cancer." Fam Cancer **10**(3): 447-453.

Sarthy, A. V., S. E. Morgan-Lappe, D. Zakula, L. Vernetti, M. Schurdak, J. C. Packer, M. G. Anderson, S. Shirasawa, T. Sasazuki and S. W. Fesik (2007). "Survivin depletion preferentially reduces the survival of activated K-Ras-transformed cells." Mol Cancer Ther **6**(1): 269-276.

Sasco, A. J., M. B. Secretan and K. Straif (2004). "Tobacco smoking and cancer: a brief review of recent epidemiological evidence." Lung Cancer **45 Suppl 2**: S3-9.

Sato, M., J. E. Larsen, W. Lee, H. Sun, D. S. Shames, M. P. Dalvi, R. D. Ramirez, H. Tang, J. M. DiMaio, B. Gao, Y. Xie, Wistuba, II, A. F. Gazdar, J. W. Shay and J. D. Minna (2013). "Human lung epithelial cells progressed to malignancy through specific oncogenic manipulations." Mol Cancer Res **11**(6): 638-650.

Savoy, R. M. and P. M. Ghosh (2013). "The changing roles of steroid nuclear receptors with prostate cancer progression." Endocr Relat Cancer **20**(4): C9-11.

Sawyer, P. R., R. N. Brogden, R. M. Pinder, T. M. Speight and Avery (1975). "Clotrimazole: a review of its antifungal activity and therapeutic efficacy." Drugs **9**(6): 424-447.

Scholl, C., S. Frohling, I. F. Dunn, A. C. Schinzel, D. A. Barbie, S. Y. Kim, S. J. Silver, P. Tamayo, R. C. Wadlow, S. Ramaswamy, K. Dohner, L. Bullinger, P. Sandy, J. S. Boehm, D. E. Root, T. Jacks, W. C. Hahn and D. G. Gilliland (2009). "Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells." Cell **137**(5): 821-834.

Schupp, M. and M. A. Lazar (2010). "Endogenous ligands for nuclear receptors: digging deeper." J Biol Chem **285**(52): 40409-40415.

Sekine, I., Y. Nishiwaki, R. Kakinuma, K. Kubota, F. Hojo, T. Matsumoto, H. Ohmatsu, M. Yokozaki, K. Goto, T. Miyamoto, J. Takafuji and T. Kodama (1997). "Phase II study of high-dose dexamethasone-based association in acute and delayed high-dose cisplatin-induced emesis--JCOG study 9413." Br J Cancer **76**(1): 90-92.

Sell, S. (2004). "Stem cell origin of cancer and differentiation therapy." Crit Rev Oncol Hematol **51**(1): 1-28.

Semizarov, D., L. Frost, A. Sarthy, P. Kroeger, D. N. Halbert and S. W. Fesik (2003). "Specificity of short interfering RNA determined through gene expression signatures." Proc Natl Acad Sci U S A **100**(11): 6347-6352.

Shang, Y. and M. Brown (2002). "Molecular determinants for the tissue specificity of SERMs." Science **295**(5564): 2465-2468.

Shao, D. D., A. Tsherniak, S. Gopal, B. A. Weir, P. Tamayo, N. Stransky, S. E. Schumacher, T. I. Zack, R. Beroukhim, L. A. Garraway, A. A. Margolin, D. E. Root, W. C. Hahn and J. P. Mesirov (2013). "ATARiS: computational quantification of gene suppression phenotypes from multisample RNAi screens." Genome Res **23**(4): 665-678.

Shay, J. W. and W. E. Wright (2011). "Role of telomeres and telomerase in cancer." Semin Cancer Biol **21**(6): 349-353.

Shiau, A. K., D. Barstad, J. T. Radek, M. J. Meyers, K. W. Nettles, B. S. Katzenellenbogen, J. A. Katzenellenbogen, D. A. Agard and G. L. Greene (2002). "Structural characterization of a subtype-selective ligand reveals a novel mode of estrogen receptor antagonism." Nat Struct Biol **9**(5): 359-364.

Shih, T. and C. Lindley (2006). "Bevacizumab: an angiogenesis inhibitor for the treatment of solid malignancies." Clin Ther **28**(11): 1779-1802.

Shopland, D. R., H. J. Eyre and T. F. Pechacek (1991). "Smoking-attributable cancer mortality in 1991: is lung cancer now the leading cause of death among smokers in the United States?" J Natl Cancer Inst **83**(16): 1142-1148.

Siegel, R., D. Naishadham and A. Jemal (2013). "Cancer statistics, 2013." CA Cancer J Clin **63**(1): 11-30.

Siemiatycki, J., L. Richardson, K. Straif, B. Latreille, R. Lakhani, S. Campbell, M. C. Rousseau and P. Boffetta (2004). "Listing occupational carcinogens." Environ Health Perspect **112**(15): 1447-1459.

Sigoillot, F. D. and R. W. King (2011). "Vigilance and validation: Keys to success in RNAi screening." ACS Chem Biol **6**(1): 47-60.

Simmer, J. P., R. E. Kelly, A. G. Rinker, Jr., J. L. Scully and D. R. Evans (1990). "Mammalian carbamyl phosphate synthetase (CPS). DNA sequence and evolution of the CPS domain of the Syrian hamster multifunctional protein CAD." J Biol Chem **265**(18): 10395-10402.

Singer, O. and I. M. Verma (2008). "Applications of lentiviral vectors for shRNA delivery and transgenesis." Curr Gene Ther **8**(6): 483-488.

Singh, A. and J. Settleman (2010). "EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer." Oncogene **29**(34): 4741-4751.

Singh, K. K., L. Wachsmuth, A. E. Kulozik and N. H. Gehring (2013). "Two mammalian MAGOH genes contribute to exon junction complex composition and nonsense-mediated decay." RNA Biol **10**(8): 1291-1298.

Sioud, M. (2006). "Innate sensing of self and non-self RNAs by Toll-like receptors." Trends Mol Med **12**(4): 167-176.

Skrypnyk, N., X. Chen, W. Hu, Y. Su, S. Mont, S. Yang, M. Gangadhariah, S. Wei, J. R. Falck, J. L. Jat, R. Zent, J. H. Capdevila and A. Pozzi (2014). "PPARalpha activation can help prevent and treat non-small cell lung cancer." Cancer Res **74**(2): 621-631.

Sladek, F. M. (2011). "What are nuclear receptor ligands?" Mol Cell Endocrinol **334**(1-2): 3-13.

Smith, C. L. and B. W. O'Malley (2004). "Coregulator function: a key to understanding tissue specificity of selective receptor modulators." Endocr Rev **25**(1): 45-71.

Sommer, S. and S. A. Fuqua (2001). "Estrogen receptor and breast cancer." Semin Cancer Biol **11**(5): 339-352.

Sood, A., D. L. Barton, B. A. Bauer and C. L. Loprinzi (2007). "A critical review of complementary therapies for cancer-related fatigue." Integr Cancer Ther **6**(1): 8-13.

Steckelberg, A. L., V. Boehm, A. M. Gromadzka and N. H. Gehring (2012). "CWC22 connects pre-mRNA splicing and exon junction complex assembly." Cell Rep **2**(3): 454-461.

Subramanian, J. and R. Simon (2010). "Gene expression-based prognostic signatures in lung cancer: ready for clinical use?" J Natl Cancer Inst **102**(7): 464-474.

Sudbery, I., A. J. Enright, A. G. Fraser and I. Dunham (2010). "Systematic analysis of off-target effects in an RNAi screen reveals microRNAs affecting sensitivity to TRAIL-induced apoptosis." BMC Genomics **11**: 175.

Sun, S., J. H. Schiller and A. F. Gazdar (2007). "Lung cancer in never smokers--a different disease." Nat Rev Cancer **7**(10): 778-790.

Sunaga, N., D. S. Shames, L. Girard, M. Peyton, J. E. Larsen, H. Imai, J. Soh, M. Sato, N. Yanagitani, K. Kaira, Y. Xie, A. F. Gazdar, M. Mori and J. D. Minna (2011). "Knockdown of Oncogenic KRAS in Non-Small Cell Lung Cancers Suppresses Tumor Growth and Sensitizes Tumor Cells to Targeted Therapy." Molecular Cancer Therapeutics **10**(2): 336-346.

Tahergorabi, Z. and M. Khazaei (2012). "A review on angiogenesis and its assays." Iran J Basic Med Sci **15**(6): 1110-1126.

Tamborero, D., A. Gonzalez-Perez, C. Perez-Llamas, J. Deu-Pons, C. Kandoth, J. Reimand, M. S. Lawrence, G. Getz, G. D. Bader, L. Ding and N. Lopez-Bigas (2013). "Comprehensive identification of mutational cancer driver genes across 12 tumor types." Sci Rep **3**: 2650.

Tang, J., R. Salama, S. M. Gadgeel, F. H. Sarkar and A. Ahmad (2013). "Erlotinib resistance in lung cancer: current progress and future perspectives." Front Pharmacol **4**: 15.

Tang, Q., Y. Chen, C. Meyer, T. Geistlinger, M. Lupien, Q. Wang, T. Liu, Y. Zhang, M. Brown and X. S. Liu (2011). "A comprehensive view of nuclear receptor cancer cistromes." Cancer Res **71**(22): 6940-6947.

Tijsterman, M. and R. H. Plasterk (2004). "Dicers at RISC; the mechanism of RNAi." Cell **117**(1): 1-3.

Tilesi, F., P. Fradiani, V. Socci, D. Willems and F. Ascenzioni (2009). "Design and validation of siRNAs and shRNAs." Curr Opin Mol Ther **11**(2): 156-164.

Topalian, S. L., F. S. Hodi, J. R. Brahmer, S. N. Gettinger, D. C. Smith, D. F. McDermott, J. D. Powderly, R. D. Carvajal, J. A. Sosman, M. B. Atkins, P. D. Leming, D. R. Spigel, S. J. Antonia, L. Horn, C. G. Drake, D. M. Pardoll, L. Chen, W. H. Sharfman, R. A. Anders, J. M. Taube, T. L. McMiller, H. Xu, A. J. Korman, M. Jure-Kunkel, S. Agrawal, D. McDonald, G. D. Kollia, A. Gupta, J. M. Wigginton and M. Sznol (2012). "Safety, activity, and immune correlates of anti-PD-1 antibody in cancer." N Engl J Med **366**(26): 2443-2454.

Torres-Arzayus, M. I., J. Font de Mora, J. Yuan, F. Vazquez, R. Bronson, M. Rue, W. R. Sellers and M. Brown (2004). "High tumor incidence and activation of the PI3K/AKT pathway in transgenic mice define AIB1 as an oncogene." Cancer Cell **6**(3): 263-274.

Travis, W. D. (2011). "Pathology of lung cancer." Clin Chest Med **32**(4): 669-692.

Tschuch, C., A. Schulz, A. Pscherer, W. Werft, A. Benner, A. Hotz-Wagenblatt, L. S. Barrionuevo, P. Lichter and D. Mertens (2008). "Off-target effects of siRNA specific for GFP." BMC Mol Biol **9**: 60.

Tsubouchi, Y., H. Sano, Y. Kawahito, S. Mukai, R. Yamada, M. Kohno, K. Inoue, T. Hla and M. Kondo (2000). "Inhibition of human lung cancer cell growth by the peroxisome proliferator-activated receptor-gamma agonists through induction of apoptosis." Biochem Biophys Res Commun **270**(2): 400-405.

Tusher, V. G., R. Tibshirani and G. Chu (2001). "Significance analysis of microarrays applied to the ionizing radiation response." Proc Natl Acad Sci U S A **98**(9): 5116-5121.

Valastyan, S. and R. A. Weinberg (2011). "Tumor metastasis: molecular insights and evolving paradigms." Cell **147**(2): 275-292.

Varlotto, J. M., L. N. Medford-Davis, A. Recht, J. C. Flickinger, E. Schaefer, D. S. Zander and M. M. DeCamp (2011). "Should large cell neuroendocrine lung carcinoma be classified and treated as a small cell lung cancer or with other large cell carcinomas?" J Thorac Oncol **6**(6): 1050-1058.

Wairagu, P. M., K. H. Park, J. Kim, J. W. Choi, H. W. Kim, B. I. Yeh, S. H. Jung, S. J. Yong and Y. Jeong (2014). "Combined therapeutic potential of nuclear receptors with receptor tyrosine kinase inhibitors in lung cancer." Biochem Biophys Res Commun **447**(3): 490-495.

Wang, H., Y. Wang, E. R. Rayburn, D. L. Hill, J. J. Rinehart and R. Zhang (2007). "Dexamethasone as a chemosensitizer for breast cancer chemotherapy: potentiation of the antitumor activity of adriamycin, modulation of cytokine expression, and pharmacokinetics." Int J Oncol **30**(4): 947-953.

Wang, X., R. K. Varma, L. Beauchamp, S. Magdaleno and T. J. Sendera (2009). "Selection of hyperfunctional siRNAs with improved potency and specificity." Nucleic Acids Res **37**(22): e152.

Wang, Y., C. E. Kaiser, B. Frett and H. Y. Li (2013). "Targeting mutant KRAS for anticancer therapeutics: a review of novel small molecule modulators." J Med Chem **56**(13): 5219-5230.

Watts, J. K., G. F. Deleavey and M. J. Damha (2008). "Chemically modified siRNA: tools and applications." Drug Discov Today **13**(19-20): 842-855.

Weinberg, O. K., D. C. Marquez-Garban, M. C. Fishbein, L. Goodglick, H. J. Garban, S. M. Dubinett and R. J. Pietras (2005). "Aromatase inhibitors in human lung cancer therapy." Cancer Res **65**(24): 11287-11291.

Weinstein, I. B. and A. K. Joe (2006). "Mechanisms of disease: Oncogene addiction--a rationale for molecular targeting in cancer therapy." Nat Clin Pract Oncol **3**(8): 448-457.

With, C. M., D. L. Evers and J. T. Mason (2011). "Regulatory and ethical issues on the utilization of FFPE tissues in research." Methods Mol Biol **724**: 1-21.

Xie, Y., G. Xiao, K. R. Coombes, C. Behrens, L. M. Solis, G. Raso, L. Girard, H. S. Erickson, J. Roth, J. V. Heymach, C. Moran, K. Danenberg, J. D. Minna and Wistuba, II (2011). "Robust gene expression signature from formalin-fixed paraffin-embedded samples predicts prognosis of non-small-cell lung cancer patients." Clin Cancer Res **17**(17): 5705-5714.

Yang, D., F. Buchholz, Z. Huang, A. Goga, C. Y. Chen, F. M. Brodsky and J. M. Bishop (2002). "Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells." Proc Natl Acad Sci U S A **99**(15): 9942-9947.

Yeung, M. L., L. Houzet, V. S. Yedavalli and K. T. Jeang (2009). "A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication." J Biol Chem **284**(29): 19463-19473.

Zhang, C., B. Beckermann, G. Kallifatidis, Z. Liu, W. Rittgen, L. Edler, P. Buchler, K. M. Debatin, M. W. Buchler, H. Friess and I. Herr (2006). "Corticosteroids induce chemotherapy resistance in the majority of tumour cells from bone, brain, breast, cervix, melanoma and neuroblastoma." Int J Oncol **29**(5): 1295-1301.

Zhang, Z., P. E. Burch, A. J. Cooney, R. B. Lanz, F. A. Pereira, J. Wu, R. A. Gibbs, G. Weinstock and D. A. Wheeler (2004). "Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome." Genome Res **14**(4): 580-590.

Zhao, D., F. Zhang, W. Zhang, J. He, Y. Zhao and J. Sun (2013). "Prognostic role of hormone receptors in ovarian cancer: a systematic review and meta-analysis." Int J Gynecol Cancer **23**(1): 25-33.

Zhong, R., J. Kim, H. S. Kim, M. Kim, L. Lum, B. Levine, G. Xiao, M. A. White and Y. Xie (2014). "Computational detection and suppression of sequence-specific off-target phenotypes from whole genome RNAi screens." Nucleic Acids Res **42**(13): 8214-8222.

Zhou, H., M. Xu, Q. Huang, A. T. Gates, X. D. Zhang, J. C. Castle, E. Stec, M. Ferrer, B. Strulovici, D. J. Hazuda and A. S. Espeseth (2008). "Genome-scale RNAi screen for host factors required for HIV replication." Cell Host Microbe **4**(5): 495-504.