NUCLEAR RECEPTORS IN LUNG CANCER

APPROVED BY SUPERVISORY COMMITTEE

David J. Mangelsdorf, Ph.D. (Mentor)

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

Carole R. Mendelson, Ph.D.

Jerry W. Shay, Ph.D.

I would like to thank

Our Heavenly Father, My father, My mother in heaven, My parents-in-law, My children

(Brian and David), and especially

My Lovely wife for her constant love and support.

NUCLEAR RECEPTORS IN LUNG CANCER

by

YANGSIK JEONG

DISSERTATION / THESIS

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

March, 2007

# NUCLEAR RECEPTORS IN LUNG CANCER

Publication No. _____

Yangsik Jeong, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2007

Supervising Professors: David J. Mangelsdorf, Ph.D.

John D. Minna, M.D.

Lung Cancer is a fatal disease with new diagnoses of more than 150,000 Americans every year. Although it has a relatively well-known etiology (e.g. smoking) and has been widely researched, clinical tools and markers for early diagnosis, prognostic prediction, and therapeutic interventions remain limited. Here, for the first time, I propose a novel translational approach for providing diagnostic, prognostic, mechanistic, and therapeutic information by studying of the expression of the nuclear receptor (NR) superfamily in lung cancer. Using quantitative real-time PCR, mRNA expression levels for the 48 members of the NR superfamily were profiled in 56 lung cell lines. Based on the resulting dataset, further analysis was performed to show the diagnostic and therapeutic potential of the NR profile using both an *in vitro* cell response assay and an *in vivo* mouse xenograft model with cognate ligand treatment for selected nuclear receptors. In addition, the NR profiles of 30

microdissected and pair-matched patient tissue samples provided a subset of NRs showing dramatic differences in expression and subgroupings that demonstrate individual variations between the normal and corresponding tumor. Furthermore, I identified several individual NRs as well as a subgroup of NRs with prognostic power. The relevance of NRs to disease pathogenesis was then studied in genetically manipulated human bronchial epithelial cells (HBEC3) and in transgenic $K\text{-}ras^{V12}$ mice, a well-known genetic model for lung adenocarcinoma. In the HBEC3 panel, the induced expression of peroxisome proliferator activating receptor gamma (PPARγ) in the parental HBEC3 introduced by oncogenic $K\text{-}ras^{V12}$ is decreased in a subset of tumorigenic clones derived from the parental cells. It appears to be strongly correlated to the expression of cylooxygenase 2 (COX2), which is shown to be decreased with PPARγ ligand treatment. In the transgenic model, I demonstrated that expression of a subgroup of NRs in wild type mice becomes altered in histologically normal tissues that harbor the K-ras mutation, and become further altered in tumor tissues of the mutant. This observation suggests that NR profiling also provides a valuable tool for understanding disease pathogenesis in lung cancer.

TABLE OF CONTENTS

Prior Publications

Bookout, A.L., **Jeong, Y**., Downes, M., Yu, R. T., Evans, R.M., and Mangelsdorf, D.J. (2006). Anatomical Profiling of Nuclear Receptor Expression Reveals a Hierarchical Transcriptional Network. Cell 126, 789-799.

Chang-Qing Xie, Mingui Fu, Minerva T. Garcia-Barrio, Zichao Zhang, Tingwan Sun, Angie Bookout, **Yangsik Jeong**, Jifeng Zhang, Ren-He Xu, Yuqing E. Chen and David J. Mangelsdorf (2007). Expression Profiling of Nuclear Receptors in Human and Mouse Embryonic Stem Cells. Molecular Endocrinology  (in press)

# LIST OF FIGURES AND TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| 18S | 18S Ribosomal RNA |
| ACTH | Adrenocorticotropic hormone |
| ADK | Adenocarcinoma |
| AF | Activation Function |
| ANOVA | Analysis of Variance |
| APC2 | adenomatosis polyposis coli 2 |
| AR | Androgen Receptor |
| BA | Bile Acid |
| BA | Bronchioalveolar carcinoma |
| BASC | Bronchial alveolar stem cells |
| BAT | Brown Adipose Tissue |
| CAR | Constitutive Androstane receptor |
| CDK | Cyclin dependent Kinase |
| C. elegans | Caenorhabditis elegans |
| CETP | Cholesteryl Ester Transfer Protein |
| CGH | Comparative Genomic Hydrization |
| COUP-TF | Chicken Ovalbumin Upstream Promoter-Transcription Factor |
| COX2 | Cyclooxygenase 2 |
| CPT | Carnitine Palmitoyltransferase |
| CREBP | cAMP Response Element-Binding Protein |
| CT | Computed Tomography |

| | |
|---|---|
| CYP7A1 | Cytochrome P450 7 alpha Hydroxylase A1 |
| DAX-1 | Dosage-sensitive Sex Reversal-adrenal Hypoplasia Congenita-critical Region on the X Chromosome Gene 1 |
| DBD | DNA binding Domain |
| Decitabine | 2'-deoxy-5-azacytidine |
| DKK1 | Dickkopf homolog 1 |
| DHT | Dihydrotestosterone |
| DNMT | DNA methyltransferase |
| DMEM | Dulbecco's Modified Eagle Medium |
| DNA | Deoxyribonucleic Acid |
| ER | Estrogen Receptor |
| EGFR | Epidermal Growth Factor Receptor |
| ERR | Estrogen-Related Receptor |
| ETS | Environmental Tobacco Smoke |
| FBS | Fetal Bovine Serum |
| FGF | Fibroblast Growth Factor |
| FXR | Farnesoid X Receptor |
| GCNF | Germ Cell Nuclear Factor |
| GPCR | G Protein Coupled Receptor |
| GR | Glucocorticoid Receptor |
| GSK3$\beta$ | Glycogen synthase kinase 3 beta |
| HAT | Histone Acetyltransferase |

| | |
|---|---|
| HBEC | Human bronchial epithelial cells |
| HDAC | Histone Deacetylase |
| HMT | Histone Methyltransferase |
| HNF | Hepatocyte Nuclear Factor |
| HprECs | human prostate epithelial cells |
| HPV | human papilloma virus |
| HR | Hazard ratio |
| hTERt | Human telomerase reverse transcriptase |
| IGF | Insulin-like growth factor |
| IGFBP | Insulin-like Growth Factor Binding Protein |
| INOS | Inducible Nitric Oxide Synthase |
| Insig | Insulin Induced Gene |
| IR | Inverted Repeat |
| IRB | Internal review board |
| LC | Large Cell Carcinoma |
| LRH | Liver Receptor Homolog |
| LOH | Loss of Heterozygosity |
| Luc | Luciferase |
| LXR | Liver X Receptor |
| LXRE | LXR Element |
| MAPK | Mitosis Activating Protein Kinase |
| ME | Mesothelioma |

| | |
|---|---|
| MR | Mineralocorticoid Receptor |
| MDR | Multi-drug Resistance |
| NcoR | Nuclear Receptor Co-Repressor |
| NGFI-B | Nerve Growth Factor Induced Gene-B |
| NNK | 4-(methylnitrosamino)-1-(3-pyridyl)-butanone |
| NSCLC | Non-small Cell Lung Carcinoma |
| NTC | Non template control |
| PCA | Principal Component Analysis |
| PCR | Polymerase Chain Reaction |
| PDK4 | Pyruvate dehydrogenase kinase 4 |
| PET | Positron Emission Tomography |
| PML | Promyelocytic leukemia |
| PNR | Photoreceptor-specific Nuclear Receptors |
| PPAR | Peroxisome Proliferator-Activated Receptor |
| PR | Progesterone Receptor |
| PTH | Parathyroid Hormone |
| PXR | Pregnane X Receptor |
| QPCR | Quantitative PCR |
| RAR | Retinoic Acid Receptor |
| Rb | Retinoblastoma |
| RNA | Ribonucleic Acid |
| ROR | Retinoic Acid-related Orphan Receptor |

| | |
|---|---|
| RVR | Reverse-erb Receptor |
| RXR | Retinoid X Receptor |
| SCA | Statistical Coupling Analysis |
| SCC | Squamous Cell Carcinoma |
| SCD | Stearoyl CoA desaturase |
| SCLC | Small Cell Lung Carcinoma |
| SF-1 | Steroidogenic Factor-1 |
| SHP | Small Heterodimer Partner |
| SMRT | Silencing Mediator of Retinoic Acid and Thyroid Hormone Receptors |
| SOM | Self-organizing map |
| SP-C | Sulfactant Protein C |
| SXR | Steroid Xenobiotic Receptor |
| T3 | Thyroid Hormone |
| T4 | Thyroxine |
| TBS | Tris-Buffered Saline |
| TG | Triglycerides |
| TGFβ | Tumor growth factor beta |
| TGFβR | Tumor growth factor beta receptor |
| TKI | Tyrosine kinase inhibitor |
| TLX | Tailless |
| TNF-α | Tumor Necrosis Factor Protein-alpha |

| | |
|---|---|
| TNM | Tumor size; Nodule involvement; Metastasis |
| TR2 / 4 | Testicular Receptors 2 / 4 |
| TRα / β | Thyroid Hormone Receptor alpha / beta |
| TZD | Thiazolidinediones |
| VDR | Vitamin D Receptor |
| WAT | White Adipose Tissue |
| WT | Wildtype |

# CHAPTER ONE
## Introduction

### 1.1 Lung Cancer

The term 'Cancer' is believed to originate from the Greek word meaning *crab*. The first use of the words 'carcino' and 'carcinoma' is credited to the Greek physician Hippocrates (460 ~ 370 B.C.), well known as the 'Father of Medicine'. Hippocrates was the first to recognize the difference between benign and malignant tumors (http://www.cancer.org, ; http://www.rare-cancer.org). Although cancer has a very long history, it remains one of the major causes of disease-related deaths, even at present. Amazingly, one out of two men and one out of three women will be diagnosed with cancer over their lifetime in the United States (ACS, 2006). Since the discovery of the disease, surgery has been the primary course of action despite the development of newer therapies, i.e. chemotherapy, radiotherapy, and most recently, molecular targeted therapy. Approximately 15 organ sites dominate 80 % of all cancers. Tumors with the highest incidence include prostate cancer (33% for men) and breast cancer (31 % for women). However, both cancer types also show the highest five-year survival. On the other hand, overall mortality is highest for other types of cancers (13 to 15 % of five-year survival for both genders), while the disease incidence is relatively low compared to prostate and breast cancers.

Lung cancer-related deaths are estimated at 160,000/year including 90,000 men and 71,000 women in the United States, 2006. A similar number of new patients are diagnosed every year. This statistic accounts for 31% of men and 26% of women in cancer-related deaths, making lung cancer the leading cause of cancer death which is secondly ranked next to deaths related to heart failure in the United States (ACS, 2006; Society, 2006). Although the campaigns to quit smoking have been effective in reducing the incidence rate in men to 70 per 100,000, the rate is still increasing in women (40 per 100,000) (ACS, 2006). Along with smoking cessation efforts, the survival improvement has doubled over the last 30 years, due to several combined treatment strategies including surgery, radiotherapy, and chemotherapy. However, the five year survival rate of this disease is very low (~15%), compared to breast (75 - 88%) and prostate cancers (67 - 100 %) in the last 3 decades (ACS, 2006). One of reasons for the high survival rate in both breast and prostate cancer is due to early detection. Therefore, to improve outcome for other cancers, new tools and molecular markers for early diagnosis are required together with the development of therapeutic, as well as chemoprevention strategies.

### 1.1.1   *Classification and Etiology*

Lung cancers develop in the respiratory epithelium, which includes bronchi, bronchioles, and alveoli. Distinct from mesothelioma and sarcoma (stromal tumors), lung tumors are divided into small cell lung carcinoma (SCLC, 18 %) originating from neuroendocrine cells and non-SCLC (73 %) both with epithelial cell origins (Minna, 2005).

The NSCLC is further subtyped histologically into adenocarcinoma, squamous cell carcinoma (SCC), large cell carcinoma and mixed types of tumors.

It is believed that this malignancy is directly or indirectly related to smoking (through environmental tobacco smoke, ETS) in more than 85 percent of patients with all types of lung neoplasms (Gazdar, 2003). In general, it is said that the relative risk in developing lung cancer is 13-fold higher in active smokers and 1.5-fold higher in indirect smokers compared to non-smokers. The term 'cigarette pack-year' represents the total amount of cigarettes in packs smoked per year and has been associated with lung cancer incidence rates. For instance, the risk for developing lung cancer is increased by 60 to 70-fold in the case of an active smoker smoking 2 packs of cigarettes a day for 20 years, compared with a non-smoker (Minna, 2005). Interestingly, women have 1.5-fold higher risk of developing lung cancer than men per given cigarette pack-year, implicating a gender difference in the susceptibility to the disease. Also, more than 50 % of lung tumors occur in ex-smokers, suggesting that lung cancer risk never returns to the non-smoker level even though the chance of the disease development is clearly decreased (Minna, 2005). This suggests that preneoplastic changes may be caused by nicotine and its derivative or smoke relevant carcinogens, which further lead to potential genetic or epigenetic alterations. Although human lung cancer is not genetically inherited, various studies characterizing this fatal disease support some acquisition of genetic or epigenetic abnormality attributed to smoking. It is speculated that chronic exposure to smoke-related carcinogens generates genetic or epigenetic changes through a multistep process that leads to the development of various lesions. The genetic abnormalities involve loss of tumor suppressors or oncogenic activation. Allelic loss of

several chromosomal regions including 3p21.3, 5q, 9p, 11q23.2 and 17p (p53) are known to be involved in lung cancer incidence. Extensive molecular studies identified individual tumor suppressor genes in the 3p21.3 chromosomal region of which loss of heterozygosity are observed in more than 90 % of both SCLC and NSCLCs. Alterations in cell cycle regulatory pathways such as G1 to S phase checkpoint (i.e. Rb-CDK4-Cyclin D1-p16 pathway) and a checkpoint of DNA damage (i.e. p53) are also known to be significantly involved in disease progression. For example, mutations in both tumor suppressor genes, RB and p53, occur in > 90% in SCLC, while p53 mutations occur in > 50% and RB mutations are found in > 20% in NSCLC. Also, the activation of oncogenic *K-ras* is observed in more than 30 % of NSCLC, especially in adenocarcinoma, whereas it is found in less than 1% of SCLC. The importance of this oncogene has been shown in mouse genetic models expressing an activated form of *K-ras* constitutively or inducibly, which autonomously develops lung adenocarcinoma. Other dominant oncogenic abnormalities include genetic amplification, rearrangement and transcriptional activation of the myc family. Overexpression of the anti-apoptotic regulator bcl-2, the growth signaling network *Her2*, and induced telomerase activity are significantly involved in lung cancer incidence. Moreover, mutations of the ras gene and c-myc amplification are associated with poor prognosis in NSCLC and SCLC, respectively. Overall, allelic loss of 3p prior to the occurrence of a 9p mutation represent the earliest events in hyperplasia which further proceed to carcinoma *in situ* and invasive cancer with p53 and ras mutations. Smoke or its related carcinogens induce epigenetic modification which is another way of regulating the expression of critical genes (i.e. tumor suppressors) or groups of genes. The epigenetic changes include non-inheritable genetic modifications, which have short-term

effects (i.e. posttranslational modifications on histones) and relatively long term effects (i.e. promoter methylation) on expression of genes critical for proliferation. Promoter hypermethylation occurs on the cytidine residues in CpG islands in the promoter region. Evidence for hypermethylation in lung cancer has been shown in the promoter regions of several genes with tumor suppressor function such as p16$^{INK4a}$ (Kim et al., 2005a; Kim et al., 2006), RASSF1A (Kim et al., 2003), RARβ (Virmani et al., 2000) and FHIT (Kim et al., 2006; Maruyama et al., 2004; Song et al., 2004). Thus, a number of pharmacological research has focused on developing inhibitors to antagonize DNA methyltransferase I (DNMTI) which would allow the re-expression of the potential tumor suppressor genes (Garcia-Manero et al., 2006; Suzuki et al., 2004). Decitabine (called 2'-deoxy-5-azacytidine) has been used in the clinics for late stage cancer patients (de Vos, 2005). Along with modulating methylation in the promoter region, research has focused on posttranslational modifications of histone proteins because histone remodeling is a prerequisite for transcriptional regulation. Santos-Rosa et al., recently summarized a large number of studies demonstrating that changes in expression, changes in activity, and mutation of histone modifying enzymes are strongly related to specific types of cancers(Santos-Rosa and Caldas, 2005). This regulation is controlled by histone modifications through acetylation by histone acetyl transferase and histone deacetylase (HDAC), phosphorylation by histone kinase, methylation by histone methyltransferase (HMT), ubiquitination, and sumoylation (Marmorstein, 2004; Santos-Rosa and Caldas, 2005; Shi and Gozani, 2005). Interestingly, the loss or reduction of CBP (a subtype of HAT acting as a coactivator) from in-frame deletion, homozygous deletion, or missense mutations is observed in lung cancer (Kishimoto et al., 2005). A promising recent

phase I/II trial using combinations of two different epigenetic modulators provides a clinically important potential in future therapeutic or chemopreventive approaches (Garcia-Manero et al., 2006; Suzuki et al., 2004).

Along with genetic or epigenetic alterations, systemic changes such as hormonal dysregulation or changes in the expression of cell surface receptors by lung cancer cells also contribute to tumor proliferation. These include peptide hormones, gastrin-releasing peptide, neuropeptides (e.g. adrenocorticotropic hormone (ACTH), parathyroid hormone (PTH), AVP, calcitonin); autocrine loop, SCF/KIT, NDF/ERBB2; receptor mediated cell growth signals, opioid receptors, nicotine receptors, and EGF receptors (Minna, 2005). Together, these data suggest that the accumulation of genetic alterations or epigenetic modification by smoke and/or carcinogens is clearly causative in lung cancer. Further progression to invasive cancer occurs with changes in the tumor microenvironment and by the loss of hormonal balance.

*1.1.2 Diagnosis and Determining Tumor Stage*

Early detection is crucial in the treatment of lung cancer. The clinical tools for early diagnosis include the screening of persons at high risk (e.g. persons with high cigarette-pack year), followed by sign or symptom screening which, if any, must be accompanied by tissue diagnosis. Various methods are employed to obtain tumor tissues including bronchoscopy for bronchial biopsy, medianoscopy for node biopsy, fine-needle aspiration for (extra) thoracic tumor biopsy guided by computed tomography scan, and more. Next, the pathologist determines the diagnosis for the type of the malignancy and its subtype, which is, in the case

of NSCLC, further staged on the basis of documented information examined according to TNM (Tumor size; Nodule involvement; Metastasis) International Staging System for lung cancer (Table1.1). For SCLC, a simple two-stage system is available: 1) limited-stage disease for SCLC limited to one hemithorax and regional lymph nodes, 2) extensive stage disease for SCLC is beyond these boundaries. Regarding general procedures for staging, chest radiographs and CT scans are performed to determine tumor size and mediastinal nodal involvement, which should be histologically documented, after complete historical and physical examination. Positron emission tomography scans are performed to detect distant metastatic disease. After tumor stage, various physiologic symptoms are considered for, so called, 'physiologic staging' which assesses a patient's conditions to determine their tolerability to various anti-tumor treatments. The physiologic considerations provide the best option for treatment methods, resectability and operability, together with anatomical staging.

In addition to the various clinical techniques in use for the early and accurate diagnosis of lung cancer, molecular approaches based on high-throughput analysis have become highlighted tools for translational research. For example, comparative genomic hybridization (CGH) arrays are available for identifying genomic rearrangements, (i.e. gene amplification, loss of heterozygosity in tumor samples. Global analysis of genetic signature changes using microarray analysis has become popular for targeting diagnostic and prognostic markers as well as for the therapeutic purpose of drug mining (Chen et al., 2007). In addition to whole genome wide screening, the customized arrays are also available that are specially designed to detect a subset of genes relevant to specific signaling pathways, i.e., kinase family (kinome), transcriptional family, or G protein-coupled receptor (GPCR). Along

with detection of genomic alterations and transcriptional changes, proteomic approaches have been taken using blood samples and sputum from patients to screen for protein markers or proteins with posttranslational modifications. Furthermore, proteomics has evolved into tissue microarray in which a series of microdissected tissue sections are mounted on plates for the purpose of simultaneous immunohistochemistry for several known markers. The combined utilization of technologies and various molecular tools is expected to improve clinical consequences. However, several limitations still remain. The technological tools such as CT and PET are highly sensitive, but have a high false-positive rate (up to 40 % in case of CT scan). The increased error rate may be overcome with advanced molecular technologies (e.g. microarray), which can be further cross-confirmed with tissue microarray. In the clinical use of microarray data, biostatistics need to be improved to deal with the huge amount of data generated so that a subset of genes (i.e., metagene or principal component set of genes) can be extracted and applied for clinical purposes (diagnosis and prognosis). It has been a concern that microarray data is not as quantitative as first anticipated despite the strength that it identifies global genetic changes. Although much improved compared to early cDNA arrays, the reproducibility of microarray data is still a critical issue considering its cost for routine use. Therefore, recently, the application of quantitative real-time PCR assays in biological research has proved its potential in various respects including quantitation, reproducibility, sensitivity, selectivity, and high-throughput capacity. Therefore, it is predicted that the future diagnostic ability will be faster and more accurate through the development of various prognostic markers using molecular approaches.

*1.1.3 Treatment Choice*

Upon diagnosis and staging of the lung tumor, treatment direction is determined. Due

to the lack of sophisticated targeted therapy, surgical resection is the basic treatment most

frequently used in various stages of NSCLCs; resection alone for early and localized tumors

(IA, IB, IIA, IIB and some IIIA), combined treatment with neoadjuvant chemotherapy for

some IIIA stages of disease (i.e., IIIA with minimal N2 involvement), combined with

preoperative or postoperative radiotherapy for stage IIIA with T3 tumors. Primary choices

for advanced IIIA (with evident N2 involvement) or some IIIB tumors are radiotherapy

and/or chemotherapy with surgical resection for IIIA with advanced N2 involvement. In

cases of stage IV and more advanced IIIB tumors, the treatment has more focus on pain and

symptom relief using radiotherapy to local sites, chemotherapy or malignant pleural effusion.

On the other hand, for SCLC, which is generally widespread, treatment is primarily by

chemotherapy with or without radiotherapy due to the lack of operability. The

chemotherapeutic drugs are devised mostly in combination selected from generic cytotoxic

drugs that have been widely used in other types of cancers and known to target DNA

replication and microtubule formation. The DNA-targeting agents include cisplatin (CDDP;

cross-linking to DNA), doxorubicin (adriamycin; inhibiting topoisomerase II), gemcitabine

(Gemzar®; interrupting DNA replication as cytidine analog), etoposide (VP-16®; acting on

topoisomerase II), irinotecan (Camptosar®; acting topoisomerase I) and pemetrexed

(Alimta®/Eli Lilly; inhibiting folate-dependent enzymes and blocking synthesis of thymidine

and nucleotides). The drugs targeting microtubule formation include: vinorelbine

(Navelbine®; binding to the tubulin of mitotic microtubules), paclitaxel (Taxol®) and

docetaxel (Taxotere®; stabilizing polymerized microtubules) and peloruside A

(RTA301/Reata) which has a similar mechanism to paclitaxel. Recent clinical outcomes

using tyrosine kinase inhibitors targeting epidermal growth factor receptor (EGFR) reflect

the importance of targeted therapy in lung cancer. Those inhibitors are gefitinib

(Iressa®/AstraZeneca) and Erlotinib (Tarceva®/Genentech) which specifically antagonize

the kinase activity of mutant EGFR. In addition, the monoclonal antibody cetuximab

(Erbitux®) which targets EGFR is currently being tested in the clinic in combination with the

kinase inhibitors or the cytotoxic drugs. Various combined treatment protocols are effective

at the beginning but become ineffective in time due to the proliferation of unresponsive

tumors. It is interesting to note that the same treatment often results in drastically different

responsiveness between individuals, suggesting that customized treatment may be needed.

This resistance could be attributed to genes involved in multi-drug resistance (i.e., MDR),

metabolism or pharmacology of the xenobiotics (i.e., further activate mechanism of

xenobiotic metabolism in the refractory tumor), or mutation in the target molecule involved

in tumors (i.e., second mutation in the mutant EGFR). The clinical observations in treatment

definitely raise the issue for development of new target(s) or new therapeutic paradigms such

as targeting residual cancer stem cells, if any, which are believed to be resistant tumors to the

treatment due to the induced differentiation. Wouldn't it be a promising therapeutic or

chemopreventive paradigm if a network of genes as a whole could be controlled? Also, isn't

it possible for customized treatment? Recent cancer research has actively focused on

diagnostic markers, which are multiple groups of genes rather than single genes, identified

from microarrays using bioinformatics tools (Chen et al., 2007; Potti et al., 2006). Similarly, the clinical trials to treat cancers are targeting multiple genes in combination.

Several epidemiological studies have shed light on the potential of chemopreventive intervention for various types of cancers including lung cancer (Sinha et al., 2003). For example, curcumin, an ingredient in spicy curry, is thought to reduce the incidence of colon, prostate, and lung cancers in India (Aggarwal et al., 2005; Shishodia et al., 2003; Sinha et al., 2003). Similarly, resveratrol, a phytoestrogen in red wine, has been shown to have anti-tumorigenic effects in cancer cell culture and mouse mammary tumor models chemically induced with a phorbol ester (Kopp, 1998; Whitsett et al., 2006). In addition, retinoic acid has been clinically applied as a chemopreventive agent in various types of cancers (Abu et al., 2005; Benner et al., 1995).

*1.1.4 Mouse model*

The therapeutic potential of new drugs is routinely assessed using cancer cell lines, a useful screening system with high-throughput capacity. However, cell lines fail to predict systemic effects (e.g., pharmacokinetics) of such treatments. Thus, xenograft models using immunocompromised mice, SCID and nude, have been preferentially used in many of the lung cancer studies to represent a more physiologically relevant model. A carcinogen induced lung cancer mouse model has also been developed using NNK (4-(methylnitrosamino)-1-(3-pyridyl)-butanone), a combination of highly tumorigenic nicotine derivatives. Although the NNK model is useful for understanding disease pathogenesis, it is

less useful for the evaluation of newly developed drugs because the mice develop mostly lung adenomas with rare cases of adenocarcinoma *in situ*, but never proceed to either NSCLC or SCLC, the most common types of human lung cancers. Since oncogenes or tumor suppressor mutations are involved in many lung cancers, genetic mouse models have been successfully developed based on these mutations for both types of lung tumors, NSCLC and SCLC. Lung tumors are autonomously developed by constitutive expression or inducible expression of oncogenic *K-ras* where the stop codon in a loxP-stop-loxP cassette is removed by *cre* recombinase expression. Meuwissen et al., generated an SCLC mouse model using cutting-edge genetic engineering technology where a "floxed" system is utilized for the conditional loss of both the Rb and p53 tumor suppressors through the addition of an adenovirus expressing cre (Ad-Cre) intrabronchially delivered. The phenotype shows remarkable similarity to human SCLC with regard to histology, immunohistochemical and metastatic behavior, and neuroendocrine features. To utilize the models as preclinical evaluation systems, it is important to know how the molecular signatures of the mouse models compare to human tumors with respect to disease progression. Molecular signatures crucial for disease progression using microarray analysis will be able to provide molecular targets for therapeutic and chemopreventive interventions.

## 1.2   Nuclear Receptors

Nuclear Receptors (NR) represent one of largest transcription factor gene families consisting of 48 members in humans, 49 members in mouse, 20 members in *Drosophila melanogaster*, and more than 280 receptors in *Caenorhabditis elegans*. The generic structure

of NRs consists of a DNA binding domain (DBD) that determines target gene specificity, a ligand binding domain (LBD) determining ligand specificity, and two activation domains, AF1 on the N-terminal part and AF2 on the C-terminal region (Figure 1.4A). Lipophilic ligands generally enter the nucleus by diffusing through the cytoplasmic and nuclear membranes, in some cases with the help of intracellular ligand binding proteins. In general, once ligand binds to the cognate receptor, a conformational change is induced that releases corepressor proteins and recruits coactivator proteins to the AF2 domain leading to transcriptional activation of target genes (Figure 1.4C). The DNA response elements, in general, are configured as direct repeats (DR), inverted repeats (IR), or everted repeats (ER) of a canonical hexad motif AGGTCA. The nomenclature is DR1 ~ DR6 or IR1 ~ IR6 based on the number of nucleotides in between the two hexanucleotide repeats. These response element configurations determine target gene specificities that are further defined when combined with tissue specific competent factors.

### 1.2.1 Classifications of Nuclear Receptor Groups

Based on discovery of physiological ligands and functions, the classification of the NR superfamily has been recently updated into the following subgroupings: (1) endocrine receptors; (2) adopted receptors; (3) orphan receptors, so named due to the lack of a known physiological ligand or functional role (Figure 1.4B). As a subset of endocrine receptors, the steroid receptors bind high-affinity (typically subnanomolar dissociation constants) endocrine ligands derived from a cholesterol backbone and are involved in maintaining various

physiologic processes: 1) androgen receptor (AR), progesterone receptor (PR) and estrogen receptors (ERα and ERβ)  for reproduction and sexual differentiation, 2) mineralocorticoid receptor (MR) for electrolyte balance, and 3) glucocorticoid receptor for carbohydrate metabolism and stress responses. Another subgroup of endocrine receptors within the steroid class includes retinoic acid receptors (α, β, γ), thyroid receptors (α, β) for body thermogenesis and the vitamin D receptor (VDR) for calcium homeostasis. Unlike the nuclear localization of most NRs, the steroid receptors but ERs are initially sequestered in the cytoplasm (and thus inactive) by binding heat shock proteins. Cognate ligand binding releases the receptors from the complex and allows for their translocation into the nucleus where the steroid receptors execute target gene activation via response elements in the promoter region. In the non-genomic mechanism of action, the cytoplasmic steroid receptors activate the MAPK kinase cascade (or vice versa), thus resulting in AP1 (jun/fos)-mediated transcriptional activation of certain genes.

The second 'adopted' class of NRs have micromolar ligand binding affinities and includes three isoforms of retinoid X receptors (RXRα, β, γ), three isoforms of peroxisome proliferator activated receptors (PPARα, β, γ), two isoforms of liver X receptors (LXRα, β), farnesoid X receptor (FXR), pregnane X receptor (PXR), and constitutive androstane receptor (CAR). The physiological functions of NRs in this class involve xenobiotic sensors in the case of CAR and PXR, lipid sensors (i.e., fatty acid sensors) in the case of PPARs and cholesterol/bile acid sensors in the case of LXRs and FXR. However, the RXRs, as obligatory partner molecules for the other receptors, play an important role in contributing 'permissivity' to the partner receptors, which will be further discussed in section 1.2.4.3. IN

spite of constant adoption of the orphan nuclear receptors by understanding their physiology, half of the mammalian receptors still remain orphans. The field is even less advanced in lower vertebrates. Only two receptors in Drosophila (ecdysone and E75 receptor) and, more recently, one receptor in C. elegans (DAF-12) have known ligands (de Rosny et al., 2006; Motola et al., 2006). The third class of orphan receptors include the remaining nuclear receptors for which physiological functions have not been clearly elucidated, although some potential ligands have been proposed for certain receptors, e.g., synthetic steroids for ERRs ($\alpha$, $\beta$ and $\gamma$), fatty acids for HNF-4 ($\alpha$ and $\gamma$), fatty acid and sterols for RORs ($\alpha$, $\beta$ and $\gamma$), and phospholipids for SF-1 and LRH-1. Clearly, research will continue for novel physiological functions and ligand identification, along with their potential coordination with higher-level physiological networks which will be further discussed in section 1.2.3.

*1.2.2 Paradigm Shifts in the Nuclear Receptor Field:*

   *From Molecular Endocrinology through Reverse Endocrinology, toward*

   *Transcriptional Physiology*

   Since the first evidence of the existence of a steroid hormone receptor was shown using radioactive estradiol, two decades passed before the GR was cloned in 1986. Further studies demonstrated the mechanism of action of lipophilic endocrine molecules is performed through receptor molecule residing in the nucleus. Newly developed techniques -- low stringency hybridization and domain swap, cell-based cotransfection assays, domain mapping -- successfully identified several nuclear receptors including TR$\alpha$, TR$\beta$, and MR, and further connected molecular biology to endocrinology, opening the 'Molecular

Endocrinology' era. Therefore, more orphan receptors were cloned since the first orphan estrogen-related receptor (Giguere et al., 1988). Adoption of the orphan retinoid X receptor (RXR) occurred upon discovery of 9-cis retinoic acid as a physiologic ligand and clearly demonstrated that these receptors would drive a new field of 'ligand discovery', leading to a new era of 'Reverse Endocrinology'(Heyman et al., 1992; Mangelsdorf et al., 1992). Furthermore, the discovery of the RXR heterodimers and a core hexad motif 'AGGTCA' for the DNA binding domain (DBD) provided a clearer picture of how nuclear receptors act on the transcriptional regulation of downstream targets as well as an index to confirm the downstream targets. Thus, most receptors (about 40 nuclear receptors at that time) were reclassified according to the configuration rules (i.e., Direct Repeat, Inverted Repeat, and Everted Repeat) mentioned in section 1.2. The continuous adoption process of the orphans (i.e., PPARs, LXR, FXR, PXR, and CAR) further prospered reverse endocrinology and even expanded the field to various types of physiologies including fat, cholesterol/bile acid, xenobiotics, and diseases relevant to those physiologies. In addition, the configuration rule sorted the previously known target genes into groups under the transcriptional control of the same NR. This classification method further applied to newly identified target genes, leading to better understanding the molecular mechanisms for various physiologic pathways. In addition, coregulators, including chromatin remodeling complexes and epigenetic regulators (e.g., HDACs and HATs), tie up NRs into a bundle of " Transcriptional Physiology" with NRs at the core. Thus, the question arises if it is possible to consider the NR superfamily as a whole for understanding the hierarchy of physiological regulatory networks at tissue or organismal scales.

*1.2.3 Higher-level Functional Networks within the NR Superfamily*

Much has been learned regarding physiological roles of certain NRs in particular types of tissues by identifying physiological ligands and downstream target genes. However, the communication between multiple receptors involved in a single or multiple tissues still remains to be understood. This complexity was elegantly addressed using a systems biology approach to obtain quantitative expression profiles of the NR superfamily in 39 tissues in two different mouse models, 129sv and C57BL6. Surprisingly, the analysis revealed two main clusters of NRs, reproduction and nutrient metabolism, suggesting that the hierarchical NR networks are coordinated in physiologic paradigms (Figure 1.7). In addition to the pattern of spatial expression, the temporal NR blueprint as a whole was mapped in four metabolically important tissues including muscle, liver, white adipose tissue and brown adipose tissue (Yang et al., 2006). Twenty-five of forty-nine mouse nuclear receptors display diurnal expression patterns. More interestingly, both thyroid receptor $\alpha$ and $\beta$ follow dramatic circadian patterns of expression while thyroid hormones are generally maintained at a constant level (Yang et al., 2006). This observation strongly supports that more receptors are potentially involved in physiologic changes related to the circadian cycle (Yang et al., 2006). Collectively, it is clear that the so called first-dimensional understanding of physiology (e.g., lining up 'receptor-ligand-target genes' in a particular tissue type) has evolved to a second-dimensional paradigm (e.g., fat regulation, in various tissues such as muscle, liver and adipose tissues), which suggests higher-order molecular coordination that includes 'time', thus leading to the complexity of body homeostasis. One elegant but simple way to

understand the megacoordination of multiple variations (for example, here nuclear receptors in space and time) is to utilize biostatistics and bioinformatics tools which provide an insight to the functional roles of the orphan receptors, as well as the coordinated roles of the clustered NRs under the same physiologic pathways or between different but relevant physiologies for harmonious regulation.

*1.2.4 NRs in Physiology and Disease*

Since extensive research toward ligand discovery has adopted several orphan NRs, many distinct physiologic pathways have been elucidated. Included in those pathways are fat, cholesterol/bile acid, immune response, differentiation, development, circadian rhythm and xenobiotic metabolism. Two representative but distinctive physiologies – fat and cholesterol/bile acid metabolism – will be described in more detail in the following sections 1.2.4.1 and 1.2.4.2.

*1.2.4.1 Fat Sensors: PPARs*

The PPAR isoforms ($\alpha$, $\beta$ and $\gamma$) exhibit distinct patterns of expression in different tissues, suggesting distinctive roles in different locations of the body for fatty acid metabolism (Rosen and Spiegelman, 2000). PPAR$\gamma$, the first PPAR discovered in this subcategory, was shown to be required for development of adipose tissue in mice (Barak et al., 1999; Rosen and Spiegelman, 2000). In addition, PPAR$\gamma$ activation is sufficient to stimulate differentiation of fibroblasts into mature adipocytes (Tontonoz et al., 1994). The physiological role of this adipogenic receptor is believed to coordinate the storage of

triglyceride (TG) in adipose tissue, preventing accumulation of TG in peripheral tissues. Tissue-specific deletion of the receptor in muscle and adipose tissue results in insulin resistance while the liver-specific knock-out develops hyperinsulinemia (He et al., 2003; Hevener et al., 2003; Norris et al., 2003). The PPARγ ligand thiazolidinedione (TZD) increases insulin sensitivity, which further reduces hepatic gluconeogenesis and induces glucose uptake in muscle. In addition, the receptor-mediated modulation of adipokine secretion and adiponectin production from adipose tissue is also thought to contribute to the receptor's insulin sensitizing function. Recent studies of the anti-inflammatory function of PPARγ will be discussed in chapter 4.

PPARα is mainly expressed in liver and binds fibrates. PPARα responds to fasting by promoting fatty acid oxidation and production of ketone bodies to provide energy to peripheral tissues. PPARα null mice display a fatty liver phenotype with chronic high-fat feeding and hepatic lipid accumulation with overnight fasting which is accompanied by hypoglycemia and elevated serum free fatty acid levels (Kersten et al., 1999). Thus, fibrate has been used as an agent for hypertriglyceridemia treatment.

Unlike the previous two PPARs (α and γ), PPARδ shows ubiquitous expression, suggesting more a systemic regulation of fatty acid homeostasis in the body. In fact, the function of PPARδ was shown to be relevant to high-fat diet-induced obesity in both tissue-specific transgenic and tissue-specific null mouse models of the receptor. PPARδ activation promotes fatty acid oxidation in muscle and thermogenesis in brown adipose tissue. The tissue-specific function of PPARδ has been elucidated for heart. Mice in which PPARδ has been specifically deleted in cardiomyocytes develop cardiomyopathy (Cheng et al., 2004).

On the other hand, the so called 'marathon mouse' which has constitutive expression of PPARδ in skeletal muscle displays enhanced running endurance (Wang et al., 2004b). This enhanced PPARδ activity transformed skeletal muscle fiber composition from 'fast-twitch' muscle which has lower numbers of mitochondria and depends mostly on glucose as the major energy source, to 'slow-twitch' muscle, which is fatigue-resistant with more mitochondria and thus burns more fat as a major energy source. Collectively, the pharmacology will be of certain benefit for various high fat-related diseases such as heart disease, diabetes, metabolic syndrome, and even severe obesity.

*1.2.4.2 Cholesterol and Bile Acid Sensors: LXR and FXR, the Yin and Yang*

LXRα, LXRβ, and FXR, members in the adopted class, play crucial physiological roles in cholesterol and bile acid homeostasis. The work was initiated with identification of oxysterols (e.g., 22(R)-hydroxycholesterol, 24(S)-hydroxycholesterol, 24(S), 25-epoxycholesterol and 27-hydroxycholesterol) as physiological ligands of LXRs (Janowski et al., 1996; Lehmann et al., 1997). Further studies using genetic mouse models demonstrated that LXRα null and liver-specific knockout mice (data not published) show a severe phenotype of hepatic cholesteryl ester accumulation with high cholesterol feeding (Peet et al., 1998a; Peet et al., 1998b). On the other hand, LXR activation reduces total body cholesterol content by decreasing cholesterol absorption, enhancing reverse cholesterol transport, and increasing cholesterol catabolism into bile acids. A series of genes involved in this pathway has been identified.  Contributing to cholesterol efflux and absorption, several ATP-cassette (ABC) binding transporters, including ABCA1, ABCG5, and ABCG8, are involved in the

reverse cholesterol transport mechanism to get rid of body cholesterol. ABCA1 plays a role in removal of peripheral lipid, cholesterol, and phospholipids from peripheral macrophages through apoA1-mediated efflux (Costet et al., 2000; Repa et al., 2000). The LXR-mediated expression of ABCG5 and 8 in the intestine lowers dietary cholesterol absorption, which contributes to lower body cholesterol (Berge et al., 2000; Repa et al., 2002). In addition to ABC transporters, several other genes relevant to decreasing body cholesterol are LXR target genes including ApoE in macrophages and adipose tissues, LPL in liver and macrophages, and human CETP (Laffitte et al., 2001; Luo and Tall, 2000; Zhang et al., 2001). In the molecular mechanism of cholesterol catabolism into bile acid (BA), LXR induces the expression of CYP7A1 (cytochrome P450 7alpha-hydroxylase), the rate-limiting enzyme in the classical pathway for bile acid production. Overproduced bile acids bind to the cognate receptor FXR, leading to SHP expression and reciprocal repression of CYP7A1 expression by complexing with the competence factor LRH-1 (Goodwin et al., 2000; Lu et al., 2000). However, the SHP-mediated repression mechanism still remains to be elucidated.

BAs are important biodetergents that solubilize dietary lipids for absorption. The enterohepatic circulation system which involves liver, gallbladder, and small intestine, controls release time and reabsorption of bile acids; greater than 90% of bile acids are reabsorbed from small intestine into gallbladder. Recent studies provide information about an additional physiological regulation of fecal bile acid secretion between the enterohepatic networks. Upon feeding, diet-mediated CCK secretion from the small intestine induces bile acid secretion into small intestine where it activates FXR directly, which in turn induces fibroblast growth factor 15 (FGF15, an ortholog of human FGF19) expression. FGF15

originating from the intestine turns out to be involved in two aspects of bile acid regulation. First, it represses CYP7A1 expression via SHP, which becomes activated through phosphorylation by the FGF15 bound FGF receptor (FGFR4) in liver (Inagaki et al., 2005). Additionally, secreted FGF15 reaches and induces refilling the shrunken gallbladder to stop BA secretion into the intestine, a novel physiologic observation opposing CCK action (Choi et al., 2006). Mice lacking FGF15 display several phenotypes: increased hepatic CYP7A1 expression and fecal bile acid secretion, and gallbladder emptying. Administration of recombinant FGF15 into null mice reversed the phenotypes to reduce CYP7A1 expression and restore gallbladder volume to the wild type level.

*1.2.4.3 Phantom Ligand Effect: Rexinoid Receptors Have Permissivity*

While the activity of NRs functioning as monomers and homodimers is primarily determined by the cognate ligand alone, RXR heterodimers exhibit their transcriptional activities with RXR bound to Rexinoid in a combined manner. Originally two papers independently reported the observation that the heterodimer complex containing RXR is able to activate target gene promoters of the partner receptors in the presence of the synthetic RXR ligand alone. This induces a conformational change of the second receptor and results in the dissociation of corepressors and recruitment of coactivators, a phenomenon termed the 'phantom ligand effect' (Schulman et al., 1997; Willy and Mangelsdorf, 1997). RXR heterodimers displays allosteric coupling within the complex, which involves three modes of activation: permissive, conditional, and non-permissive. Permissivity can be defined as a distinct feature of the partner NR, where the ligand-bound RXR activates genes with the

cognate response elements of the partner receptor. There is one representative example for each mode described (Shulman et al., 2004). In the first mode, LXR/RXR heterodimers display permissivity for both ligands, meaning that the complex can be activated by each receptor's ligand alone or both ligands in a more than additive manner.

In the non-permissive mode, RXR ligand is not involved in partner VDR activity with or without VDR ligand. Although the RXR ligand is not necessary, heterodimerization is necessary for VDR activation in this mode. In conditional permissivity, there is dependence on the presence of the partner receptor ligand to achieve full responses. In this example, the RXR ligand is involved in the activity of the complex to a lesser extent than permissive, but more than non-permissive action. To identify critical amino acid residues responsible for the allosteric communication, statistical coupling analysis (SCA) was developed. SCA is a method to assess if two amino acid residues have thermodynamically coevolved in two different positions. It uses statistical clustering analysis and a large scale of screening mutagenesis (Lockless and Ranganathan, 1999). In fact, amino acid residues important for permissivity were identified in both FXR and PPARα, nuclear receptors for which RXR exhibits the permissive mode of action. Mutation of the specific residues results in non-permissivity for FXR and conditional permissivity for PPARα (Shulman et al., 2004). It has been proposed that the binding partners of RXR follow the conserved rule for recognizing potential response elements. Taken together, definition of these distinct mechanisms will lead to better understanding of finely tuned genetic control and allow development of drugs differentiating allosteric modes.

1.3 Cancer and Nuclear Receptors

Since the first demonstration that TR$\alpha$ (also known as c-erbA due to its homology with the viral oncogene v-erb) is related to cancer, more than 23,000 studies have been published regarding NR biology in cancer in the last two decades (Sap et al., 1986; Weinberger et al., 1986). Considering that NRs are involved in crucial physiological pathways such as response to diet, reproduction, immune defense and development, it is not surprising that almost every receptor, including certain orphans, is potentially relevant to various types of cancers. A few selected NRs including the androgen receptor (AR), estrogen receptor (ER), PPAR$\gamma$, and VDR will be discussed as representative examples. These receptors are significantly relevant to 3 of the top 4 cancers: breast for ER, prostate for AR, colon for VDR and PPAR$\gamma$ with high incidence rates. Secondly, the chemotherapeutic interventions using drugs targeting those receptors are either in the clinic or in developmental stages.

1.3.1 Androgen Receptor (AR) and Prostate Cancer

Prostate cancers have the highest incidence rate in men in the United States and comprise 33 % of total United States cancer patients. Due to the existence of early detectable diagnostic maker, prostate specific antigen (PSA), which could partly contribute to such a high incidence rate, mortality is very low, with more than 90 % of five year survival rates (Figure 1.2). Treatment options for localized neoplasm is mostly dependent on surgery,

combined with irradiation. However, androgen-ablations have been the main option for prostate cancer treatments in the last 50 years for most unconfined prostate cancers, which are androgen-dependent but become independent of the hormonal therapy over time (Debes and Tindall, 2004). Androgen blockade includes flutamide, bicalutamide (casodex), nilutamide for anti-androgen, and luteinizing hormone releasing hormone (LHRH) analogs or antagonists. Although hormonal therapy works well at the beginning, the response becomes refractory due to various molecular changes classified into two main categories. First, AR mutation or amplification sensitizes tumorigenic cells to even small amounts of androgen, stimulating positive tumor growth even with anti-androgens. Mutations can also result in constitutively active forms of the AR. Second, due to acquired activation of other survival pathways such as functional loss of PTEN, overexpression of Bcl-2, or the autocrine action of neuropeptides, prostate cancers become androgen-independent. In this regard, several independent clinical publications report that combined chemotherapy, commonly including docetaxel, improved median survival as well as life quality of advanced, progressive, androgen-refractory prostate cancer patients (Chen and Petrylak, 2004; Petrylak et al., 2004; Tannock et al., 2004). These clinical data suggest that treatments for prostate cancers will be greatly enhanced if combined with, preceded by, or followed by AR-targeted therapies.

1.3.2 Estrogen Receptor (ER) and Breast Cancer

Breast cancers have the highest incidence rate in women, comprising more than 30% of total cancer patients in the United States. Mortality is very low due to relatively early detection and better treatment responses. Even with high five-year survival rates, the

causative factors still remain to be defined to understand the disease pathogenesis. Several genetic or epigenetic causes, e.g., myc, HER2 overexpression, p53 mutations, promoter methylation of p16$^{ARF}$, etc., have been reported but are mostly known to be not inheritable except in the case of BRCA1 and BRCA2 mutations.  Germ line mutations in these two genes are reported to increase life-time risk of developing breast cancers, ovarian cancer or both by 50 to 85 % (Casey, 1997; Marcus et al., 1997).

Diagnosis is determined on the basis of mammography and personal physical examination with symptoms, and further confirmed with tissue biopsies and histological analysis. Moreover, prognostic evaluation and treatment decisions are established on the basis of histological examination of the expression of estrogen and progesterone receptors, and more recently together with *HER2,* which is a membrane receptor for growth signals from outside cancer cells (Hortobagyi et al., 1998; Piccart-Gebhart et al., 2005). ER alpha expression occurs in 60 to 80% of primary tumors but becomes ER negative in distant, metastatic tumors. In general, there is poor prognosis for cases of breast cancer without any expression of the receptors and it becomes better in the order of expression of ER alone, ER plus PR, and both receptors together with *HER2* because of more treatment options. Tamoxifen or other selective estrogen receptor modulators (SERMs) become the first-line of hormonal treatment, in combination with progestin in case of expression of both receptors, ERα and PR. In addition, vitamin D receptor is reported to be a beneficial prognostic marker together with ERα expression in breast cancer treatment (Hussain et al., 2003; Wietzke et al., 2005). Herceptin, a monoclonal antibody against the HER2 receptor which blocks growth signaling, improves disease-free survival among advanced, HER2 positive women and is

believed to be a valuable alternative in case of refractory hormone-based therapy. Further understanding the pathogenesis using high-throughput molecular techniques, such as microarray and quantitative real-time PCR (qRT-PCR) assays in laser-captured breast tissues according to various stages of the disease, will give more targetable options. Here, expression profiling of the NR superfamily as a whole using qRT-PCR has been surveyed in a panel of 35 breast cancer cell lines. The profile was further confirmed with the histological expression of ER and PR together with the microarray dataset.

1.3.3 NRs and Colon Cancer

Colon cancer is the third most common cause of cancer-related deaths after lung cancer, prostate cancer for men, and breast cancer for women in the United States. The epidemiologic studies suggest that various factors, both environmental and genetic, are involved in colon cancer incidence. Some somatic mutations in genes involve oncogenic alterations in genes such as ras, c-myc, etc., and loss of tumor suppressors including familial adenomatous polyposis (FAP), Rb, p53, and deleted in colon cancer (DCC) (Link et al., 2005; Majumdar et al., 2004; Mitchell, 1992). In addition, dietary factors are known to be involved in the disease incidence or progression. For example, a western diet including high-fat content induces more secretion of primary bile acids which are further converted into secondary bile acids, deoxycholic acid (DCA) and lithocholic acid (LCA) which are very toxic and potential tumor promoters or co-carcinogens. Considering that BAs are physiological ligands for FXR, this receptor is expected to be relevant to colon tumorigenesis but needs to be further studied for its beneficial potential (Makishima et al., 1999; Parks et

al., 1999). As for beneficial dietary factors, calcium is relatively well-known for its protective role in colon cancer incidence and believed to be relevant to vitamin D metabolism involving VDR. VDR regulates calcium homeostasis in response to its physiological ligand 1,25α dihydroxyvitamine–D3 and is known to have a protective role in skin cancer, and possibly in colon cancer as well. Interestingly, the secondary bile acid LCA is a physiological ligand for VDR which may sense the overproduced secondary BA and prevent or bypass its toxic effects by regulating genetic signature (Cesario et al., 2006; Jalving et al., 2005; Makishima et al., 2002). Likewise, PPARγ has been demonstrated to have a protective role against tumor incidence and positive therapeutic effects on established tumors in various preclinical models (Cesario et al., 2006; Jalving et al., 2005). Considering that PPARγ is involved in modulation of anti-inflammatory pathways and non-steroidal anti-inflammatory drugs (NSAID) have a significant negative correlation to colon cancer incidence, the beneficial effect of PPARγ activation in tumorigenesis is not surprising. Overall, finding more NRs responsive to dietary factors, together with understanding their biologic relevance, is believed to provide better options for translational applications to the clinic.

1.3.4 Relevance of NRs to Lung Cancer

The relevance of GR to lung tissue was reported 10 years before its molecular cloning, although the physiological pathways had not been clearly elucidated (Ballard and Ballard, 1974; Ballard et al., 1974; Khuri and Lippman, 2000; Lippman et al., 1993; Wang et al., 2006d). Since then, various ligands, e.g, dexamethasone, for this receptor have been utilized to control unfavorable immune responses induced by transplantation or inflammation

potentially induced by chemotherapeutic treatments for the final stages of various types of cancers including lung cancers. Much has been studied for various NR involved in lung cancer. For example, treatment with the RAR ligand, all trans-retinoic acid (AtRA), has been proposed as a potential chemopreventive factor against lung cancer incidence (Khuri and Lippman, 2000; Lippman et al., 1993). In addition, treatment with the synthetic RXR ligand bexarotene was shown to reduce lung tumor progression by 50% in mouse genetic models haboring K-ras mutations (Wang et al., 2006). In response to AtRA or its derivative treatment, expression of one of the receptor subtypes, RARβ, has been shown to be a beneficial prognostic marker for lung cancer. Further studies regarding epigenetic regulation such as promoter methylation proposed that hypermethylation of one RARβ isoform, RARβ1, is involved in retinoid resistance in lung tumorigenesis (Hershberger et al., 2005; Nakagawa et al., 2005a; Nakagawa et al., 2005b; Petty et al., 2005; Stabile et al., 2005). Recently, estrogen receptors have been studied for their potential as therapeutic targets to treat lung cancer, in combination with EGFR inhibitors (Stabile et al., 2005; Hershberger et al., 2005). More recently, VDR activation by ligands has been proposed as a beneficial marker against lung cancer metastasis (Nakagawa et al., 2005a; Nakagawa et al., 2005b). Obviously, the evidence suggests that changes in various NR activities are involved in lung tumorigenesis. Thus, I performed a systematic approach to investigate expression of all members of NR superfamily in several pathologically well-defined systems.

1.4 Hypothesis and Specific Aims

Since it is evident that certain NRs are individually involved in specific types of cancers and they are also known to modulate various physiologic pathways in a coordinated fashion, I hypothesized that quantitative expression profiling of the NR superfamily as a whole would provide a novel blueprint for therapeutic targets, diagnostic and prognostic evaluation, and eventually customized treatment in lung cancer. In addition, this approach would give an insight to understanding the disease pathogenesis. Thus, I set up four specific aims to assess this hypothesis.

Aim 1. Generate NR expression profiles in a panel of lung samples

Aim 2. Perform diagnostic and prognostic evaluation of the NR superfamily in a lung cancer model

Aim 3. Validate functional aspects for the expression profile using in vitro and in vivo models

Aim 4. Identify NRs relevant to the disease pathogenesis

1.4.1 Aim 1: Generate NR Expression Profiles in a Panel of Lung Samples

To explore NR expression in various sets of lung samples, a full set of quantitative PCR primers and probes for the 48 members of the human NR superfamily were designed, validated, and used in the TaqMan® assay (HTTP://WWW.NURSA.ORG). Expression profiles of the 48 members of the family plus two isoforms, PPARδ2 and PPARγ2, were generated for these lung samples. In addition, microarray experiments were performed in a subset of lung samples to identify unique genetic signatures concordant with receptor expression. Included for this approach was a panel of lung cancer cell lines, primary human

lung tissues (tumors vs. corresponding normal tissues), a panel of immortalized human bronchial epithelial cells with or without oncogenic alterations, and the transgenic lung cancer mouse model expressing constitutively active *K-ras*$^{V12}$.

1.4.2 Aim 2: Perform Diagnostic and Prognostic Evaluation of the NR Superfamily in a Lung Cancer Model

The NR expression profile was further interrogated using several bioinformatics tools to address answers for fundamental questions relevant to diagnosis, i.e., can the NR expression profile distinguish between tumor types? Can it differentiate the normal from the tumor? Then what receptor(s) is(are) involved ?  Can NRs be used as diagnostic markers? Bioinformatics tools included unsupervised, hierarchical cluster analysis and bootstrapping for statistical analysis of clustering accuracy. In addition to diagnostic evaluation, the analysis was performed to evaluate prognostic potential for patient survival. Using log-rank test,  multivariate Cox regression analysis, and principal component analysis (called metagene analysis), the NR profiles in patient tissues were further analyzed to identify any single NR or a subset of NRs (principal component) showing predictable power. Furthermore, a subset of NRs was identified as a predictive group for patient prognosis using risk score for individual patients.

1.4.3 Aim 3: Validate Functional Aspects for the Expression Profile Using *In Vitro* and *In Vivo* Models

Along with diagnostic and prognostic potential, the expression profile was further investigated to determine if any NRs can be a therapeutic target for clinical application. Three receptors, AR, ERα and PPARγ, were chosen for preclinical study and several cell lines positively or negatively expressing each receptor were treated with cognate ligands and assayed for cell growth response. Furthermore, *in vivo* treatment of tumor was followed using a xenograft mouse model for *in vivo* evaluation of PPARγ.

1.4.4 Aim 4: Identify NRs relevant to the disease pathogenesis

NRs were surveyed in pathogenic models relevant to the tumorigenic process. The pathogenic models included a series of immortalized human bronchial epithelial cells with or without oncogenic alterations and tumorigenic clones, and normal and corresponding tumor tissues from transgenic *K-ras$^{V12}$* mouse model. Further studies with representative NRs were performed for their relevance in disease progression. The NRs important for pathogenesis could be used as targets for chemoprevention.

## 2006 Estimated Cancer Cases in the USA

| | | **Men 720,280** | **Women 679,510** | | |
|---|---|---|---|---|---|
| Prostate | 33% | | | • 31% | Breast |
| Lung & bronchus | 13% | | | • 12% | Lung & bronchus |
| Colon & rectum | 10% | | | • 11% | Colon & rectum |
| Urinary bladder | 6% | | | • 6% | Uterine corpus |
| Melanoma of skin | 5% | | | • 4% | Non-Hodgkin lymphoma |
| Non-Hodgkin lymphoma | 4% | | | • 4% | Melanoma of skin |
| Kidney | 3% | | | • 3% | Thyroid |
| Oral cavity | 3% | | | • 3% | Ovary |
| Leukemia | 3% | | | • 2% | Urinary bladder |
| Pancreas | 2% | | | • 2% | Pancreas |
| All other sites | 18% | | | • 22% | All other sites |
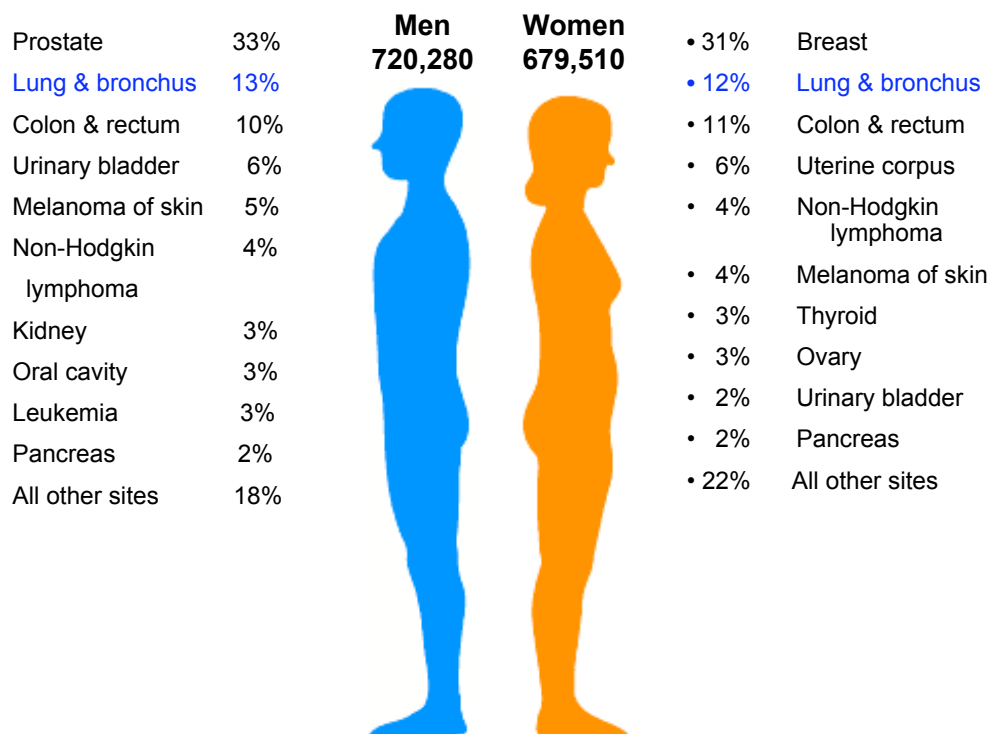
Figure 1.1 2006 Estimate of Cancer Deaths in the United States. More than Fifty percent of cancer cases are ranked into the top 4 cancers, lung, colon &rectum, prostate for men and breast for women. (Adapted by 2006, American Cancer Society, Inc)

# 2006 Estimated Cancer Deaths in the USA

| Men 291,270 | | Women 273,560 | |
|---|---|---|---|
| Lung & bronchus | 31% | • 26% | Lung & bronchus |
| Colon & rectum | 10% | • 15% | Breast |
| Prostate | 9% | • 10% | Colon & rectum |
| Pancreas | 6% | • 6% | Pancreas |
| Leukemia | 4% | • 6% | Ovary |
| Liver & intrahepatic bile duct | 4% | • 4% | Leukemia |
| Esophagus | 4% | • 3% | Non-Hodgkin lymphoma |
| | | • 3% | Uterine corpus |
| Non-Hodgkin lymphoma | 3% | • 2% | Multiple myeloma |
| Urinary bladder | 3% | • 2% | Brain/ONS |
| Kidney | 3% | • 23% | All other sites |
| All other sites | 23% | | |

**ONS=Other nervous system.**
**Source: American Cancer Society, 2006.**

Figure 1.2 2006 Estimate of Cancer Deaths in the United States. About Fifty percent of cancer deaths are attributed to the top 4 cancers, lung, colon &rectum, prostate for men and breast for women. (Adapted by 2006, American Cancer Society, Inc)

## Five-year Survival*(%) by during Three Time Periods by Cancer Site

| Site | 1974-1976 | 1983-1985 | 1995-2001 |
|---|---|---|---|
| • All sites | 50 | 53 | 65 |
| • Breast (female) | 75 | 78 | 88 |
| • Colon | 50 | 58 | 64 |
| • Leukemia | 34 | 41 | 48 |
| • Lung and bronchus | 12 | 14 | 15 |
| • Melanoma | 80 | 85 | 92 |
| • Non-Hodgkin lymphoma | 47 | 54 | 60 |
| • Ovary | 37 | 41 | 45•† |
| • Pancreas | 3 | 3 | 5 |
| • Prostate | 67 | 75 | 100 |
| • Rectum | 49 | 55 | 65 |
| • Urinary bladder | 73 | 78 | 82 |

*5-year relative survival rates based on follow up of patients through 2002.
†Recent changes in classification of ovarian cancer have affected 1995-2001 survival rates.
Source: Surveillance, Epidemiology, and End Results Program, 1975-2002, Division of Cancer Control and Population Sciences, National Cancer Institute, 2005.

Figure 1.3 Five-year Survival Rate for Three Time Periods in the United States. Breast and Prostate cancers show remarkably increased survival in the last 3 time periods, while the rest of cancers, but lung and pancreas, show mild improvement regarding survival. (Adapted and Modified by 2006, American Cancer Society, Inc)

| Stage | TNM Descriptors | 5-Year Survival Rate, % | |
|---|---|---|---|
| | | Clinical Stage | Surgical-Pathologic Stage |
| IA | T1 N0 M0 | 61 | 67 |
| IB | T2 N0 M0 | 38 | 57 |
| IIA | T1 N1 M0 | 34 | 55 |
| IIB | T2 N1 M0 | 24 | 39 |
| IIB | T3 N0 M0 | 22 | 38 |
| IIIA | T3 N1 M0 | 9 | 25 |
| | T1-2-3 N2 M0 | 13 | 23 |
| IIIB | T4 N0-1-2 M0 | 7 | <5 |
| | T1-2-3-4 N3 M0 | 3 | <3 |
| IV | Any T any N M1 | 1 | <1 |

TUMOR (T) STATUS DESCRIPTOR

| | |
|---|---|
| T0 | No evidence of a primary tumor |
| TX | Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial w hings but not visualize by imaging or bronchoscopy |
| TIS | Carcinoma in situ |
| T1 | Tumor >3 cm in greatest dimension, surrounded by lung or viceral pleura, without bronchoscopic evidence of invasion more proximal than lobar bronchus (i.e., not in main bronchus) |
| T2 | Tumor with any of following: >3 cm in greatest dimension; involves main bronchus, ≥2 cm distal to the carcina; invades visceral pleura; associated with atelectasis or obstructive pneumonitis extending to hilum but does not involve entire lung |
| T3 | Tumor of any size that directly invades any of the following: chest wall (including superior sulcus tumors), diaphragm, medistinal pleura, parietal pericardium; or tumor in main bronchus <2 cm distal to carina but without involvement of carina; or associated atelectasis or obstructive pneumonitis of entire lung |
| T4 | Tumor of any size that invades any of the following: mediastinum, heart, greatvessels, trachea, esophagus, vertebral body, carina; or tumor with a malignant pleural or pericardial effusion[a], or with satellite tumor nodule(s) within the ipsilateral primary-tumor lobe of the lung. |

TUMOR (T) STATUS DESCRIPTOR

| | |
|---|---|
| NX | Regional lymph nodes cannot be assessed |
| N0 | No regional lymph node metastasis |
| N1 | Metastasis to ipsilateral peribronchial and/or ipsilateral hilar lymph nodes, and intrapulmonary nodes involved by direct extension of the primary tumor |
| N2 | Metastasis to ipsilateral mediastinal and/or subcarinal lymph node(s) |
| N3 | Metastasis to contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s) |

TUMOR (T) STATUS DESCRIPTOR

| | |
|---|---|
| MX | Presence of distant metastasis cannot be assessed |
| M0 | No distant metastasis |
| M1 | Distant metastasis present |

Table 1.1TNM (Tumor, Node, Metastasis) International Staging System for Lung Cancer (Referenced by 2005, HARRISON's Principles of Internal Medicine, 15th Edn)

# Nuclear Receptor Superfamily

**A**

NH2- | AF-1 | DNA | Ligand | AF-2 | -COOH

**B**

## Endocrine Receptors
### Endocrine lipid sensors

| | |
|---|---|
| ER$\alpha,\beta$ | estrogen |
| AR | androgens |
| PR | progesterone |
| GR | glucocorticoid |
| MR | mineralocorticoid |

| | |
|---|---|
| RAR$\alpha,\beta,\gamma$ | retinoic acid |
| TR$\alpha,\beta$ | thyroid hormone |
| VDR | vitamin D |

## Adopted Orphan Receptors
### Dietary & endogenous lipid sensors

| | |
|---|---|
| RXR$\alpha,\beta,\gamma$ | 9 cis retinoic acid |
| PPAR$\alpha,\beta,\gamma$ | prostanoids, fatty acids |
| LXR$\alpha,\beta$ | oxysterols |
| FXR | bile acids |
| PXR | xenobiotics |
| CAR | xenobiotics |

## Orphan Receptors
### Endogenous ligands unceratin

| | |
|---|---|
| ERR$\alpha,\beta,\gamma$ | synthetic steroids |
| HNF-4$\alpha,\gamma$ | fatty acids? |
| ROR$\alpha,\beta,\gamma$ | fatty acids, sterols? |
| SF-1 | phospholipids? |
| LRH-1 | phospholipids? |
| GCNF | ? |
| PNR | ? |
| TLX | ? |
| TR 2,4 | ? |
| NGFIB$\alpha,\beta,\gamma$ | ? |
| COUP-TF$\alpha,\beta,\gamma$ | ? |
| RVR$\alpha,\beta$ | ? |
| DAX-1 | ? |
| SHP | ? |

**C**
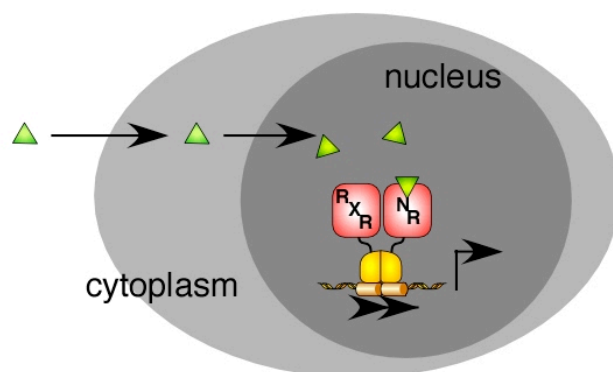
Figure 1.4 The Nuclear Receptor Superfamily
(A) Schematic representation of a typical nuclear receptor. (B) Classification of nuclear receptors according to physiological ligands. (C) Representation of NR/RXR heterodimer activation. Ligand binding to the receptor (represented as triangle) causes dissociation of corepressors and recruitment of coactivators, including transcription of a target gene. (Modified from Chawla et al., 2001)
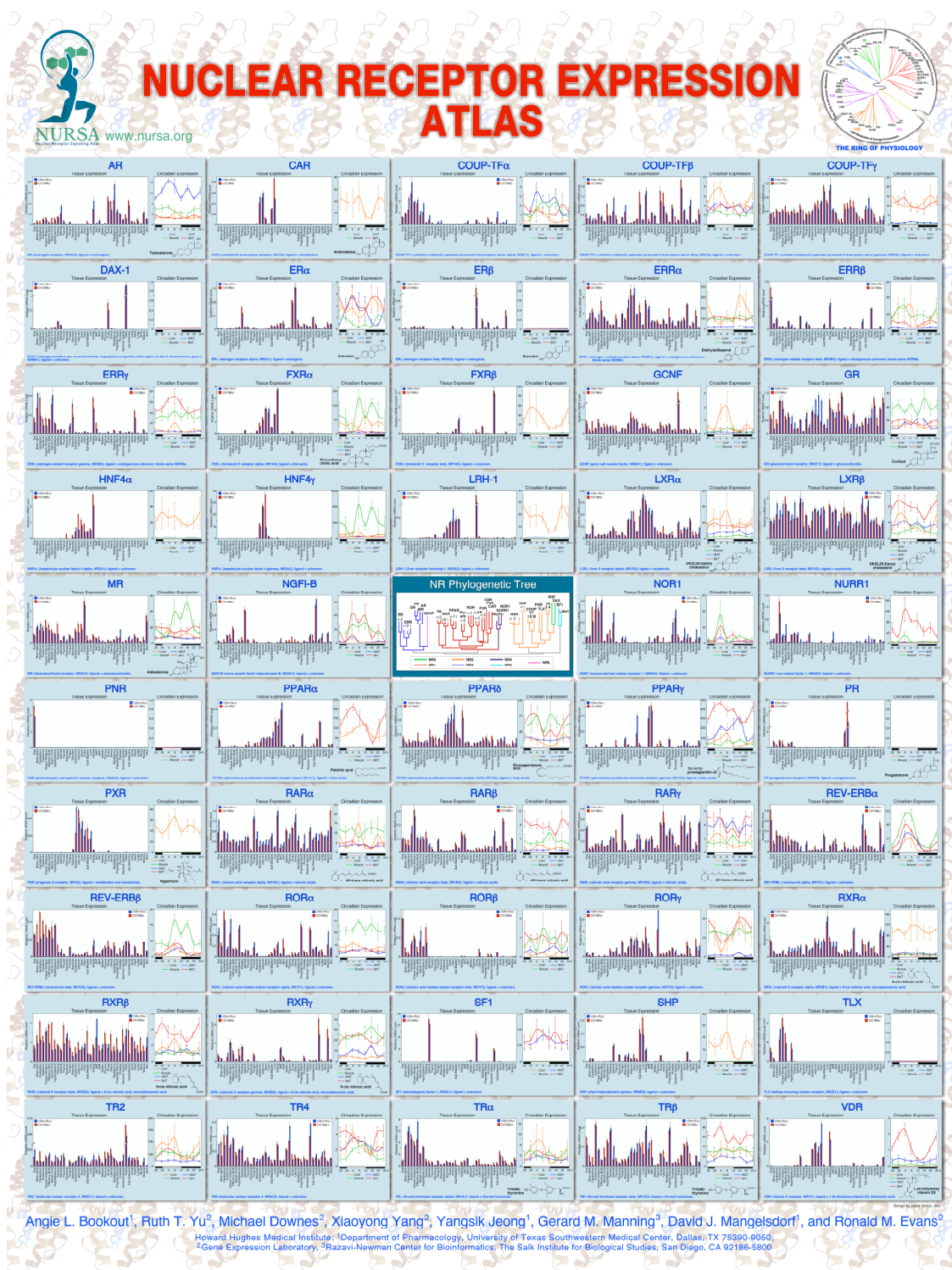
Figure 1.5 Anatomical Profiling of NR Superfamily.
(Referenced by Bookout et al, Cell 2006 and Yang et al., Cell 2006)
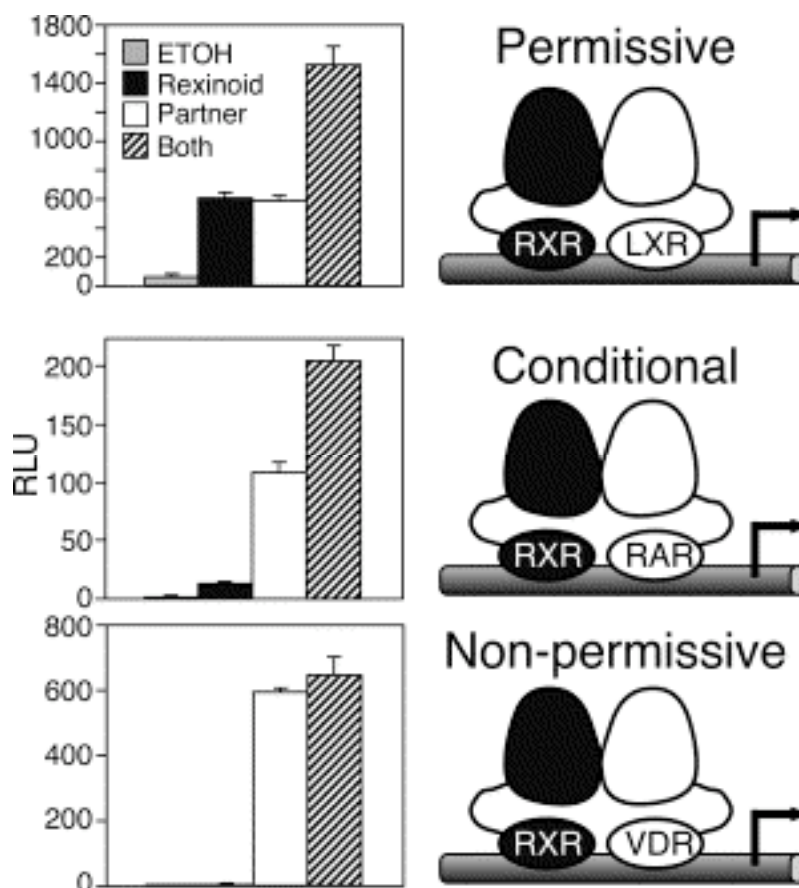
Figure 1.6 Three Activation Modes of RXR Heterodimers.
Top: Both receptors present dual permissivity in the presence of a single ligand and more than additive activation with both ligands, respectively. Middle: Conditional permissivity is allowed in the presence of RXR ligand alone which leads to full activation with RAR ligand. Bottom: The RXR ligand does not affect activation of the partner receptor. Adapted from Sulman et al, Cell 2004.
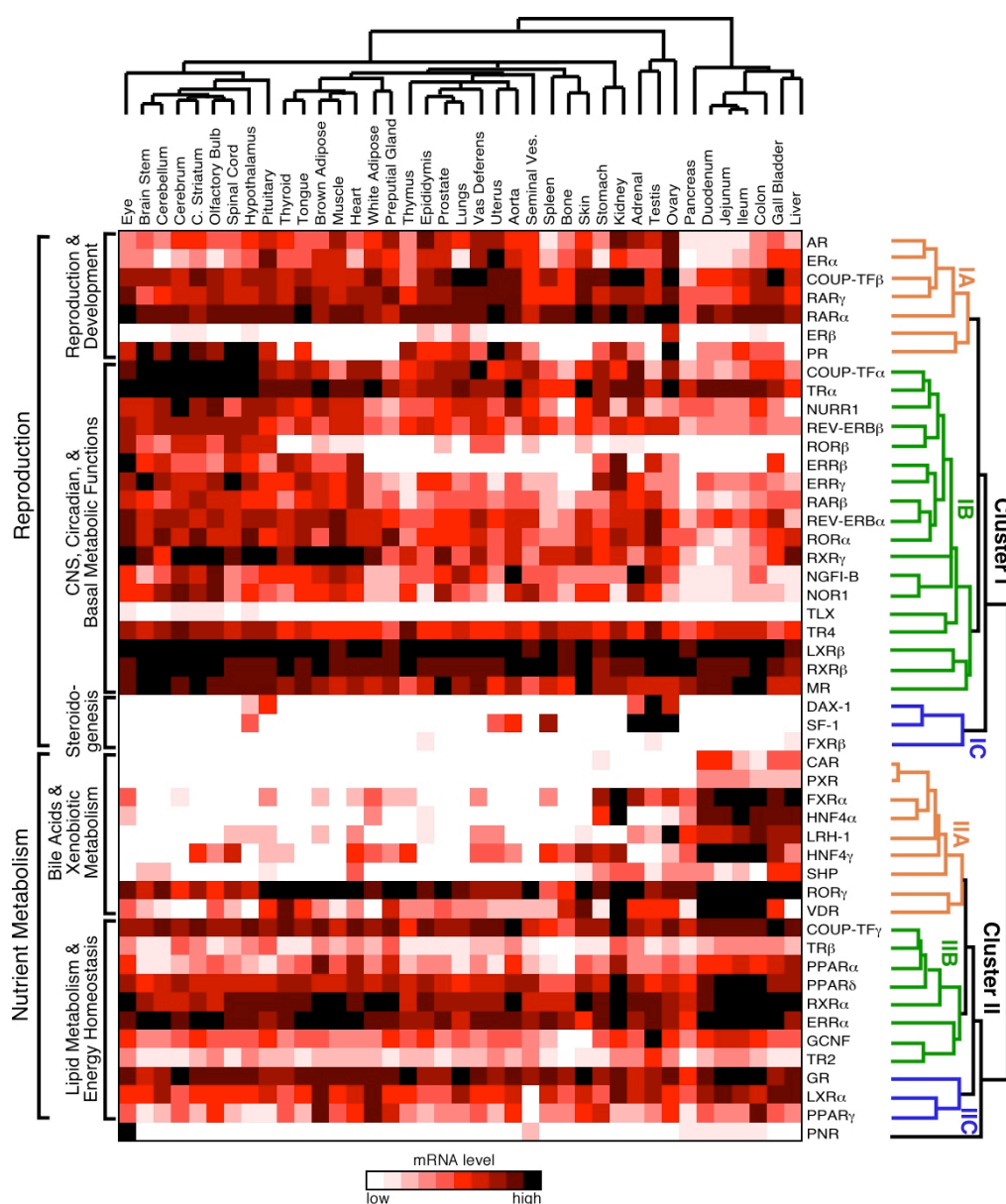
Figure 1.7 Unsupervised Hierrarchical Cluster Analysis of Mouse NR Profiles. Matrix 1.28 software was utilized to generate the hierarchical clustering results for the 49 nuclear receptor dataset that is normalized, quantitative real-time PCR values in a panel of 39 mouse tissues of two strains. The horizontal axis represents 39 mouse tissues in several subsets of clusters and the vertical axis for 49 NR sets in physiological relevant clusters. Adapted from Bookout et al, Cell 2006.

# CHAPTER TWO
## Description of Primary Data

### 2.1 Introduction

Much has been learned concerning the physiological pathways of individual NRs. The most recent work utilized a systems biology approach to assess the anatomical expression profiles of NRs as a whole family and provided a seminal insight to the existence of a higher-level regulatory network, termed ′mega-network' which maintains whole body homeostasis by balancing various physiological pathways (Bookout et al., 2006; Yang et al., 2006). This novel viewpoint has not only a "space" component, but also "time" component (Yang et al., 2006). Thus, dysregulation of the mega-network of NRs leads to imbalanced physiologies, in space and/or time, and results in many acute or chronic diseases. This new paradigm for considering the NR superfamily as a whole, rather than single receptor or dominant factor, drove me to test its potential for translational application to lung cancer as the first disease model. TaqMan-based quantitative real-time PCR assays, which are highly quantitative, reproducible, and sensitive, were employed to accomplish this goal, along with high-throughput Affymetrix microarray experiments. An array of lung samples included a panel of 56 lung cell lines and 30 pairs of patient samples (tumor and pair-matched adjacent normal tissues) for preclinical diagnostic and therapeutic evaluation, and *in vivo* prognostic evaluation of NR profiles. In addition, the NR profile was produced for two pathogenic models: first, a series of immortalized HBECs, non-tumorigenic HBECs, and tumorigenic

HBECs with oncogenic alterations; secondly, a mouse lung cancer model with oncogenic *K-ras* expression (Johnson et al., 2001; Ramirez et al., 2004; Sato et al., 2006). More detailed information about these samples are described in each of the following sections along with the primary dataset.

**2.2 Description of Samples and NR Profiles**

2.2.1 Expression of Nuclear Receptors in Lung Cell Lines

To explore the NR expression in a panel of lung cells and patient tissues, a full set of quantitative PCR primers and probes were designed and validated for the 48 members of the human NR superfamily and used in the TaqMan® assay (Bookout et al., 2006; Ramirez et al., 2004; Sato et al., 2006). Included in a panel of 56 lung cell lines were five normal human bronchial epithelial cell lines, ten immortalized cell lines, fourteen SCLCs, and twenty-seven NSCLCs. The immortalized cell lines included two BEAS2B cell lines exogenously expressing SV40 T antigen, four human bronchial epithelial cell lines (HBEC) immortalized with cyclin-dependent kinase 4 (CDK4) and the catalytic subunit of human telomerase reverse transcriptase (hTERt), three HBECs with oncogenic *K-ras$^{V12}$*, and one HBEC with human papilloma virus (HPV) E6/E7 oncoproteins, which was characterized in detail in previous studies (Ramirez et al., 2004; Sato et al., 2006). The expression profile of the 48 NRs plus two isoforms, PPARδ2 and PPARγ2, is shown in alphabetical order in figure 2.1 and further grouped by expression level into 3 main categories (Table 2.1). This analysis revealed a highly distinct but remarkably variant pattern of NR expression in different lung cancer cell lines. Some receptors show a highly predictable pattern of expression (e.g.,

NSCLC vs. SCLC), whereas others display unpredictable and sparse but distinct expression, which may be indicative of individual lung cancer differentiation or differentiating type of lung tumors. Moreover, some NRs show neither significant changes nor expression at all across the entire panel. Additionally, global genetic signatures of 48 lung cell lines out of the 56 were surveyed using affymetrix microarray chip which monitors 40,000 genes. Pearson correlation coefficients of the 48 NRs from the qRT-PCR assays were determined for all 40,000 genes on the microarray and further sorted according to the calculated Pearson coefficients. An example for VDR is represented by sorting all genes against VDR in Figure 2.2. This type of analysis revealed genetic signatures possibly under the same transcriptional control as the nuclear receptor and thus identified potential target genes for the specific NR. Considering the higher-level NR network, this bioinformatics approach may provide an insight to understanding physiological communication within a subset of NRs on how genetic signatures from individual receptors are integrated. In other words, the physiological correlation, whether positive or negative, between two nuclear receptors can be elucidated at genetic levels. Moreover, it will be interesting to apply this method to all transcription factor families and the integrative computational tools generated may help to produce virtual pathways that can be simply tested using an *in vitro* system.

2.2.2 Human Bronchial Epithelial Cells (HBECs)

Since certain NRs clearly showed differentiation of immortalized HBECs from lung cancer cell lines in section 3.2.1.1, I rationalized that these subsets of NRs are potentially involved in lung cancer disease progression. Thus, to more systemically identify NRs

potentially relevant to lung cancer pathogenesis, NR expression profiles were surveyed in a panel of immortalized HBECs previously generated by oncogenic alterations. The establishment of a series of isogenic HBECs with defined oncogenic changes is cartooned in figure 2.3. Primary bronchial epithelial cells that migrated from biopsied bronchus tissues were immortalized by CDK4 and hTERt, followed by further oncogenic manipulations. The cell lines either express *K-ras$^{v12}$*, have p53 knocked-down, or harbor both changes (Ramirez et al., 2004; Sato et al., 2006). Clonogenic methods, such as soft agar and liquid colony formation assay, isolated several clones that were further tested for their tumorigenic potential and divided into non-tumorigenic (HBEC C6, C7, C8, and C9) and tumorigenic clones (HBEC C1 and HBEC C5) using a xenograft model. The tumorigenic clones developed two distinct but typical types of lung tumors, squamous cell carcinoma for the C1 clone and adenocarcinoma for the C5 clone. In addition, these grown tumors were further reestablished as tumorigenic cell lines named C1 #658 and C5 #453. Now using these genetically defined systems, I profiled the expression of the 48 human NRs to dissect their physiological relevance in oncogene-induced pathogenesis as well as for nuclear receptors that could serve as targets for chemoprevention. Interestingly, the profile revealed distinct groups of expression, no expression, expressed but no change between clones, and differentiated expression with tumorigenecity or oncogenic alterations (Figure 2.4). Further biological studies are discussed in chapter 5.

2.2.3 Patient Tissue Samples

Along with the potential of NRs as diagnostic and therapeutic targets in preclinical *in vitro* cell culture systems (Section 2.2.1) and xenograft models (Section 4.2.4), the prognostic power of the NR superfamily was assessed using a panel of primary lung tissue samples consisting of tumor and corresponding normal tissues obtained by microdissection from a panel of patients tissue samples (n=30) that were histologically characterized into adenocarcinoma (n=24) and squamous cell carcinoma (n=6). Demographic features are described in table 2.3. The expression profile of the 48 main NRs plus two specific isoforms, PPARδ2 and PPARγ2, is shown in figure 2.5, together with NR classification according to expression level (Table 2.2). The expression profile revealed a subset of NRs showing dramatic differences of expression in tumor vs. normal tissues across all patients (Figure 2.6). Five receptors, NGFIB3, PPARγ2, PR, RXRγ, and SHP, show decreased expression in tumor vs. normal tissues, suggesting potential roles of these receptors as anti-tumorigenic factors. In contrast, HNF4γ shows increased expression in tumor vs. normal tissues. This subset of NRs can be utilized as diagnostic markers and the HNF4γ can be a potential therapeutic target for treating lung cancer using gene therapy. In addition, certain NRs clearly showed distinct patterns of expression (e.g., increased or decreased expression in tumor vs. corresponding normal tissue) in a patient-dependent manner whereas others had no difference in expression for both pair-matched tissues across all patient samples (Figure 2.7). Included in the category of individual variation are AR, COUP-TFα, COUP-TFγ, GR, MR, LRH-1, LXRβ, NOR1, NURR1, PPARγ, RARα, RARβ, RARγ, RORβ, RORγ, RXRα, TRβ and VDR. Considering NRs as druggable targets, some of which already have been clinically targeted for certain types of cancers as previously mentioned in section 1.3, the treatment decision should be

carefully chosen for selecting target NRs. In fact, retrospectively, the therapeutic responses significantly vary in generic treatments using cytotoxic drugs, and even targeted therapy for various types of cancers, which, however, had been attributed to individual variations. Here, I propose a novel therapeutic paradigm, 'tailored-drug treatment', which customizes treatment options by screening the best NR or best set of NRs based upon individual variations.

### 2.2.4 Tissue Samples from an Oncogenic *K-ras*$^{V12}$ Mouse Model

Along with multiple NRs identified in a series of isogenic HBECs as having potential pathogenic relevance to oncogenic changes, a survey of NRs was executed in an oncogenic *K-ras*$^{V12}$ mouse model, which was described in section 1.1.4. Lung tissues were obtained from groups of 4 month and 8 to 9 month old mice where each group had 10 wild types and 10 mutants including both genders. All environmental conditions, including 'diet', were controlled. Furthermore, tumor tissues were carefully dissected-out, together with separate collections of pair-matched normal tissues. Histopathologic examination of all samples in the 8 to 9 month old group revealed that tumors developed into adenocarcinoma showing typical gland formation (Figure 2.8). The NR profile of mouse lung tissues provided two interesting groups of NRs based on expression pattern. In one of these groups, 8 out of 50 NRs showed expression differences between tumor and pair-matched normal tissue in a individual mouse-specific manner (Figure 2.9). Since NRs are well-known drug targets, these data implicate the use of NR profiling as a strategy for individualized treatment against lung cancer. In the other group, there was a dramatic difference between expression of normal tissue and tumors for 10 of the 50 NRs, which may provide potential diagnostic markers as well as therapeutic

targets. This genetic model for lung cancer is believed to be an important and valuable system to assess the potential of the screened NR in therapeutic intervention as well as chemoprevention. Further bioinformatics analysis will be discussed in section 5.2.4.

2.3 Discussion

Since much has been studied concerning the significance of individual NRs in various physiologies and as potential drug targets by screening or developing ligands to modulate receptor function, here, I utilized high-throughput, quantitative real-time PCR methods to establish the expression profile of the 48 NR superfamily as a whole in various lung cancer models as the first disease target for translational research of the NR superfamily. To see if there are potential drug targets which can be used to treat, predict, and/or prevent the disease, I investigated the NR profile for various indices in several systems, such as diagnosis and therapeutic power in human lung cell lines, prognostic relevance in patient samples, and for chemoprevention in pathogenic models. The following chapters will provide more detailed discussion of the primary data using various functional, biostatistics, and bioinformatics tools.

Figure 2.1 Expression Profiles of the Nuclear Receptor Superfamily in Lung Cell Lines. Quantitative real-time PCR assay was performed for 50 NRs in 56 human lung cell lines. Relative expression values on the y-axis were obtained as described in chapter 7. The x-axis represents cell names in various types of lung cells including SCLC, small cell lung cancer; NSCLC, non-SCLC; SCC, squamous cell carcinoma; LC, large cell carcinoma; BA; ME, mesothelioma.
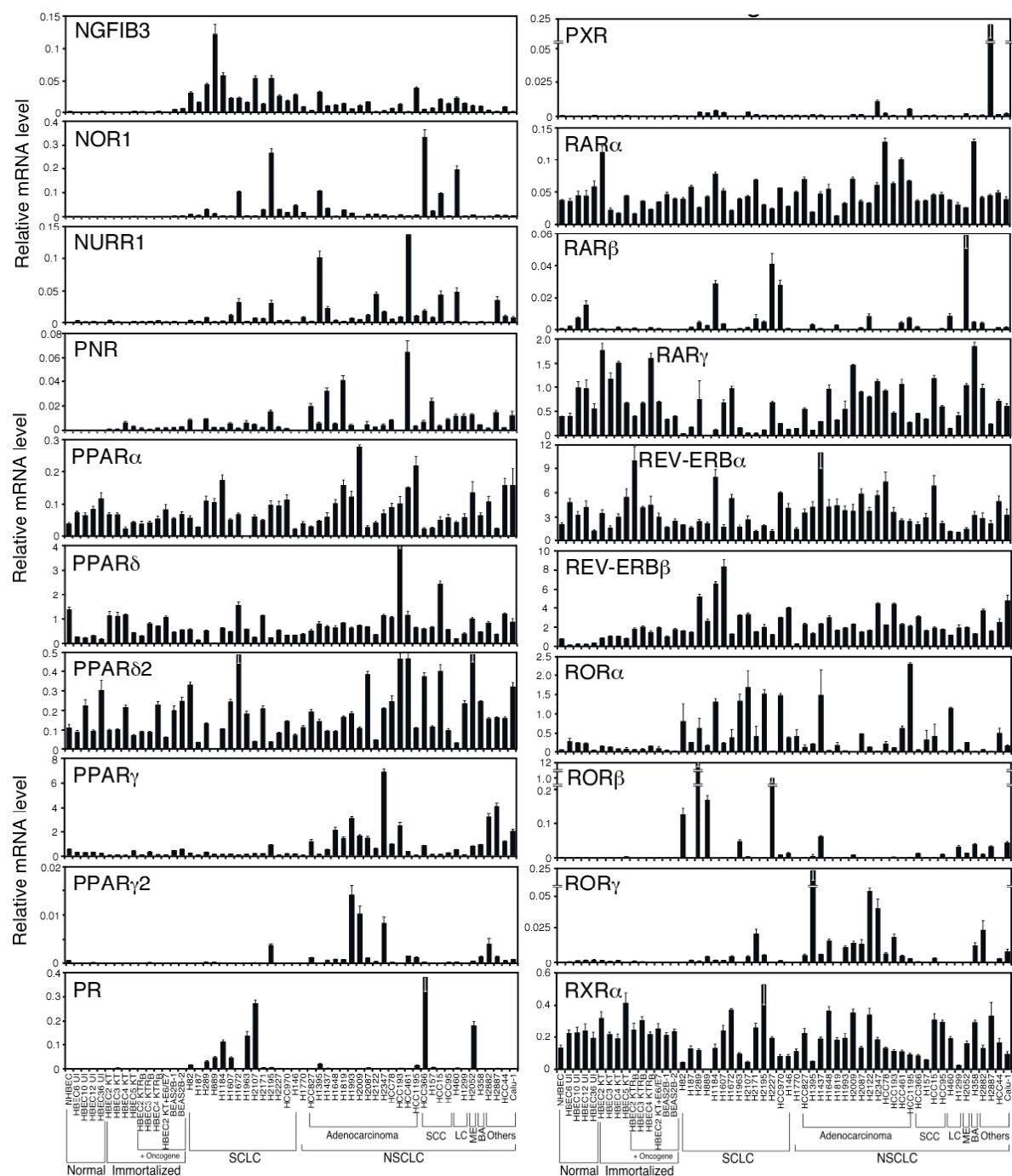
Table 2.1 Classification of NR on Expression Level

| Samples | High (expression level,≤5**) | | Intermediate (≤ 0.1) | | Low (≤ 0.01) | |
|---|---|---|---|---|---|---|
| Lung Cells | COUP-TFα | RARγ | AR | PPARα | CAR* | RARβ |
| | COUP-TFβ | REV-ERBα | DAX-1 | PR | ERβ | SF-1 |
| | COUP-TFγ | REV-ERBβ | ERα | RORγ | ERRβ | TLX |
| | GR | RORα | ERRα | SHP | ERRγ | |
| | LXRβ | RORβ | GCNF | TRβ | FXR | |
| | NURR1 | RXRα | HNF4γ | VDR | HNF4α | |
| | PPARδ | RXRβ | LXRα | | LRH-1 | |
| | PPARδ2 | RXRγ | MR | | PNR | |
| | PPARγ | TR2 | NGFIB3 | | PPARγ2 | |
| | RARα | TR4 | NOR1 | | PXR | |
| | | TRα          n=21 | | n=16 | | n=13 |

The expression profiles are grouped into high ($\sim 5$), intermediate ($0.01 \sim 0.1$), and Low (less than 0.01) where the y-axis represents *normalized value* as described in methods. The asterisk (*) represents receptors in absence.
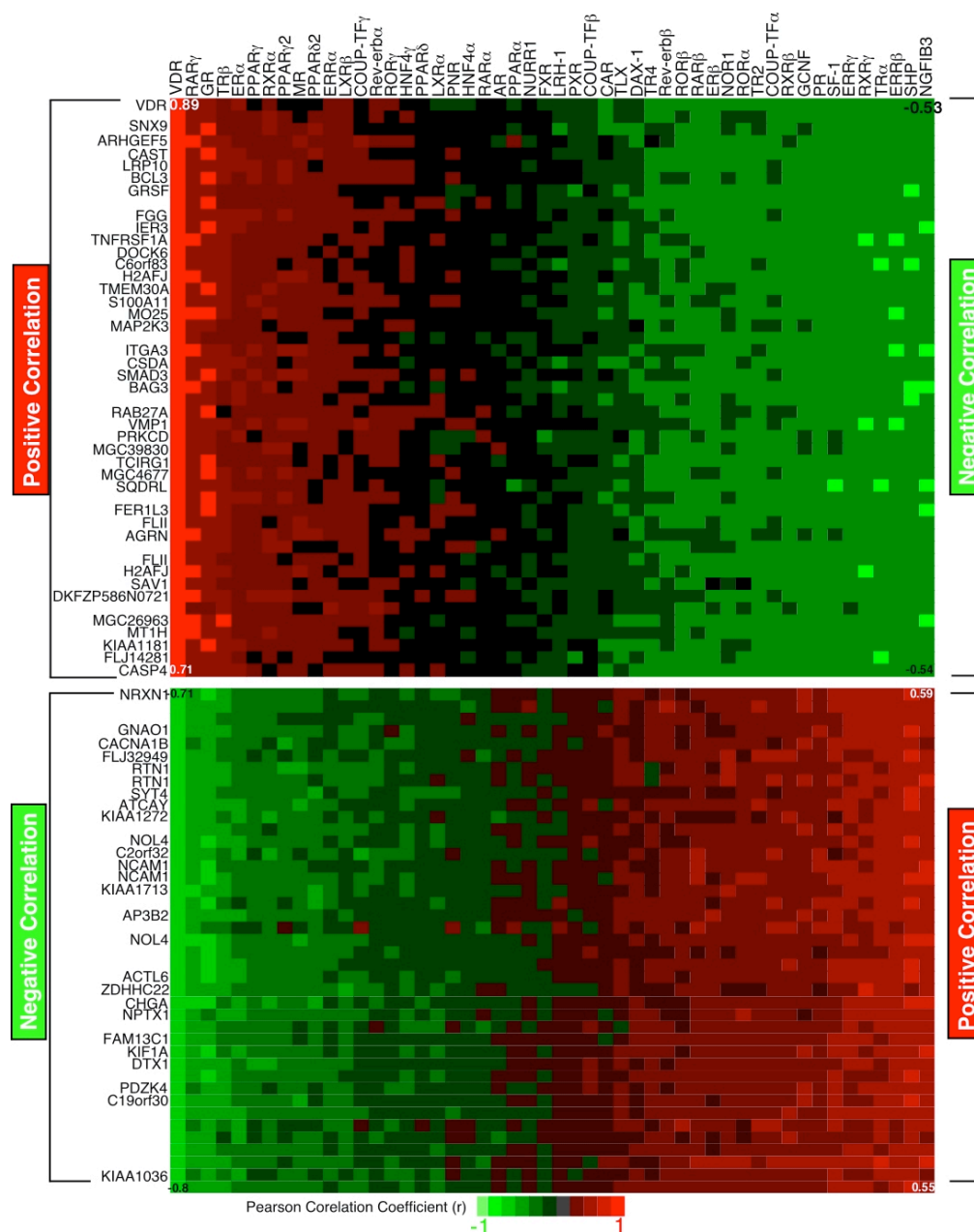
Figure 2.2. Global Change of Genetic Signatures on the VDR Expression.
Quantitative RT-PCR expression of the individual NRs is correlated to the whole gene
expressions on Affymetrix chip in 48 lung cell lines. The horizontal axis represents
NRs and the vertical axis stands for gene names on Affymetrix chip. The calculated
Pearson correlation coefficients are sorted to the VDR and represented into color codes,
red for positive and green for negative correlation values.
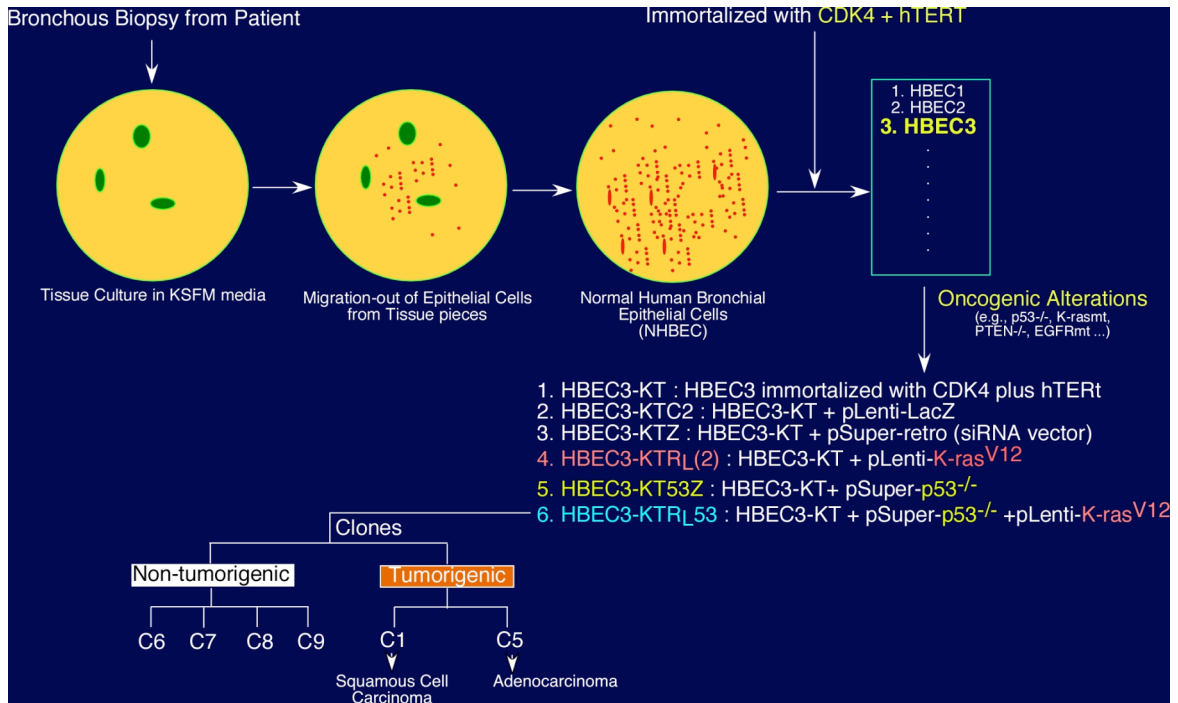
Figure 2.3. Overall Schema for Generating Immortalized and Tumorigenic Bronchial Epithelial Cells.

Human bronchial epithelial cells were collected from slice cultures of normal bronchial tissues from patients. Further genetic manipulations of HBECs include immortalization with CDK4 and hTERt and tumorigenesis with oncogenic alterations which are either introduction of oncogenic K-ras, p53 knock-down, or both. Abbreviations: CDK4, cyclin-dependent kinase 4; hTERt, human telomerase catalytic subunit; pSuper-retro, siRNA vector; plenti-, lentiviral vector.

A

**Low or no Expression**

AR, CAR, DAX, ERβ, ERRβ, ERRγ, FXR, HNF4α, HNF4γ, LRH-1, NOR1, PNR, PPARγ2, PR, PXR, RORβ, RORγ, RXRγ, SF-1, SHP, TLX

**Expressed, but no difference**

CoupTF II, EAR2, ERRα, GCNF, GR, LXRα, LXRβ, MR, NURR1, RARα, RARγ, Re-verbβ, RXRα, TRα, TRβ, TR2, TR4, VDR

1 2 3 4 5 6

1. HBEC3-KT : HBEC3 immortalized with CDK4 plus hTERt
2. HBEC3-KTC2 : HBEC3-KT + pLenti-LacZ
3. HBEC3-KTZ : HBEC3-KT + pSuper-retro (siRNA vector)
4. HBEC3-KTR$_L$(2) : HBEC3-KT + pLenti-K-ras$^{V12}$
5. HBEC3-KT53Z : HBEC3-KT+ pSuper-p53$^{-/-}$
6. HBEC3-KTR$_L$53 : HBEC3-KT + pSuper-p53$^{-/-}$ +pLenti-K-ras$^{V12}$

**Expressed differently**

CoupTF I, ERα, NGFIB3, PPARα, PPARδ, PPARδ2, PPARγ, Re-verbα, RORα, RXRβ

HBEC3 (Human Bronchial Epithelial Cell)

- Vector controls
- K-ras$^{V12}$
- p53$^{-/-}$
- p53$^{-/-}$+K-ras$^{V12}$

B

**Low or no Expression**

CAR, ERα, ERRβ, ERRγ, FXR, HNF4α, NGFIB3, PNR, PPARγ2, PR, PXR, RORγ, RXRγ, SF-1, SHP, TLX

**Expressed, but no difference**

EAR2, ERβ, GCNF, GR, LXRα, LXRβ, MR, PPARα, PPARδ, PPARδ2, RARα, RARγ, Rev-erbα, Rev-erbβ, RORα, RXRα, RXRβ, TRα, TRβ, TR2, TR4

1 2 3 4 5

HBEC3-KTR$_L$53 (Human Bronchial Epithelial Cell)

1. HBEC3-KT : HBEC3 immortalized with CDK4 plus hTERt
2. KTR$_L$ 53 : HBEC3-KT + pLenti-K-ras$^{V12}$ + pSuper-p53$^{-/-}$ in KSFM
3. C-1 : HBEC3-KTR$_L$53 Clone 1 in KSFM
4. C-5 : HBEC3-KTR$_L$53 Clone 5 in KSFM
5. C-5 tumor: HBEC3-KTR$_L$53 Clone 5 tumor

**Expressed differently**

AR, CoupTF I, CoupTF II, DAX-1, HNF4γ, LRH-1, NOR1, NURR1, RARβ, RORβ, ERRα, PPARγ, VDR

Figure 2.4. Expression Signature of the NR Superfamily in Various Pathologic Stages of HBECs. (A) Expression profiles of NRs in HBECs with different immortalization treatments. (B) Expression profiles of NRs in tumorigenic HBEC clones (C1 and C5). The profiles are classified into three groups including low or no expression, expressed but no difference, and expressed differently.
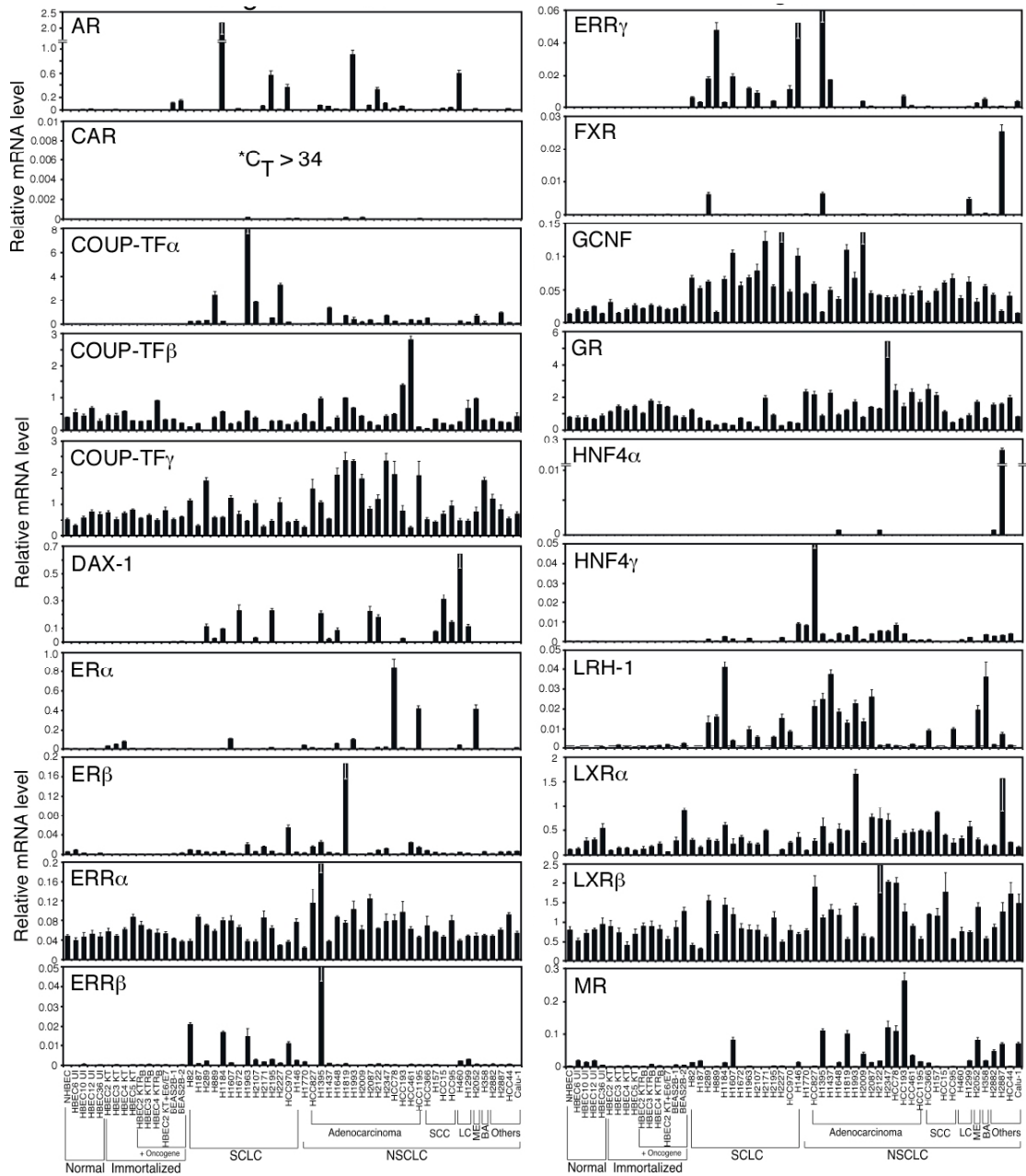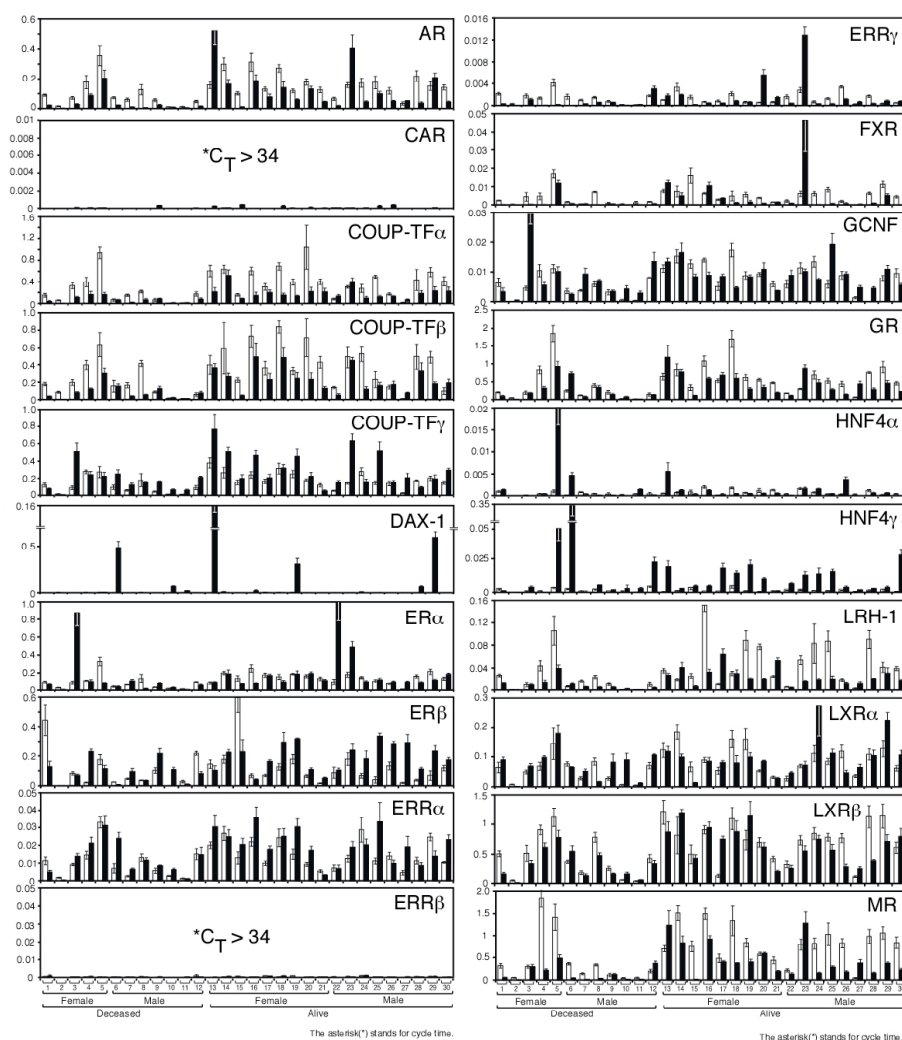


Figure 2.5 Expression Profiles of the Nuclear Receptor Superfamily in Patient Tissues. Quantitative real-time PCR assay was performed for 50 NRs in 30 pairs of human lung tissues. Relative expression values on the y-axis were obtained as described in chapter 7. The x-axis represents identification number of tissues according to the gender and in the order of survival length.

The asterisk(*) stands for cycle time.
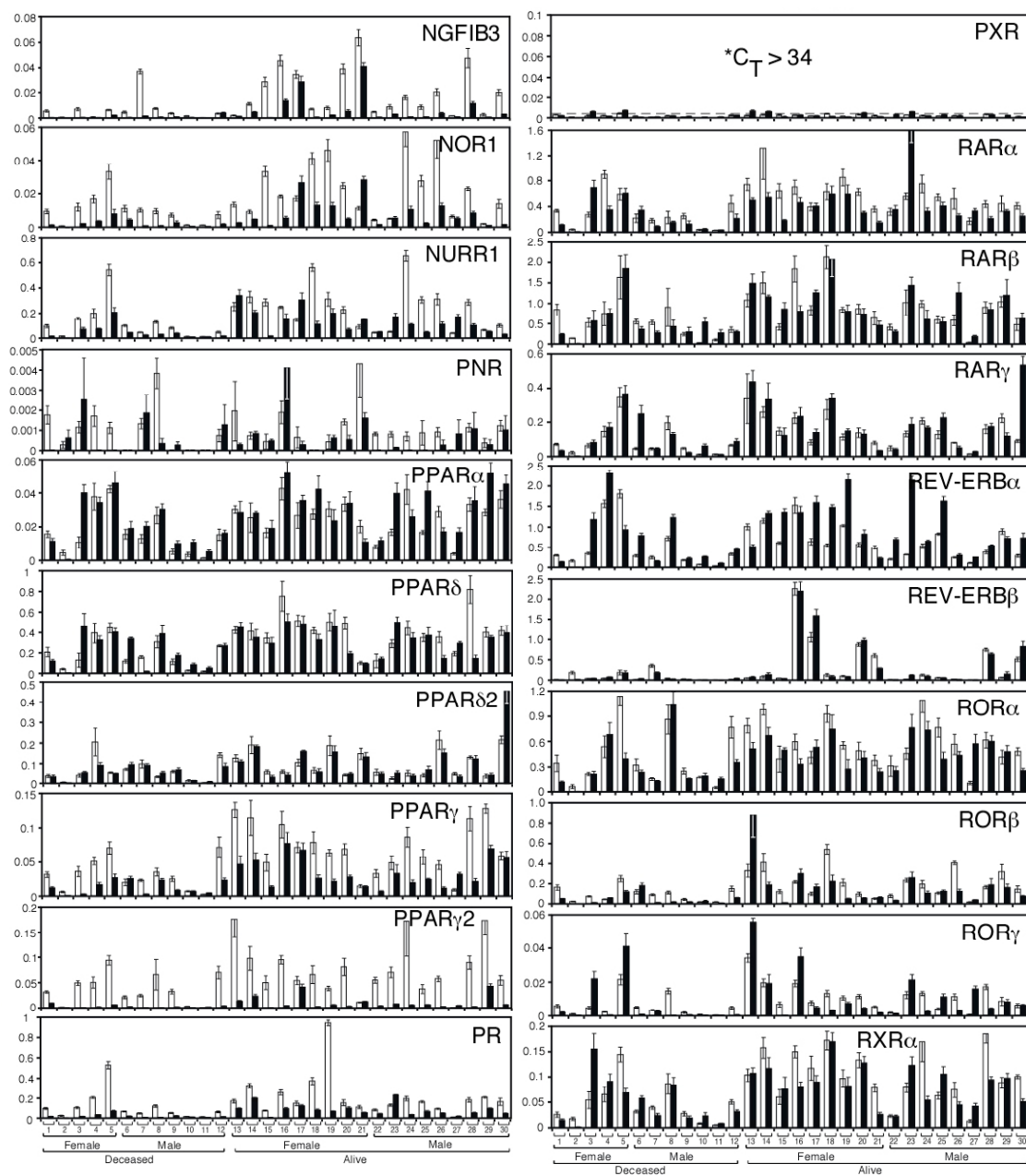
The asterisk(*) stands for cycle time.
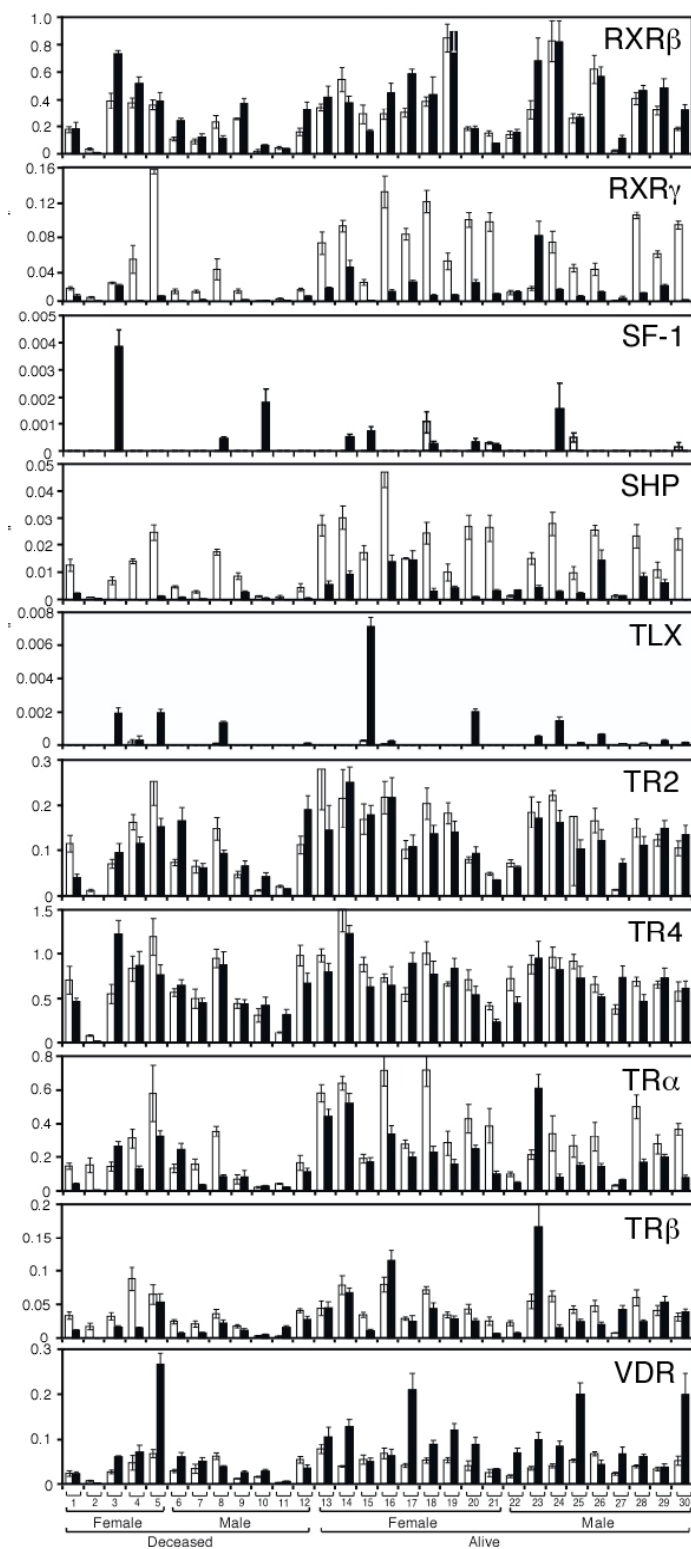
Table 2.2 Classification of NR Expression Level in Patient Samples

| | High (≤ 1) | | Intermediate (≤ 0.2) | | | Low (≤ 0.01) | |
|---|---|---|---|---|---|---|---|
| Tissues | COUP-TFα<br>COUP-TFβ<br>COUP-TFγ<br>GR<br>LXRβ<br>PPARδ<br>PPARδ2<br>PPARγ<br>RARα<br>RARγ | REV-ERBα<br>REV-ERBβ<br>RORα<br>RXRβ<br>TR4<br>TRα<br><br><br><br>n=16 | AR<br>DAX-1<br>ERα<br>ERβ<br>ERRα<br>GCNF<br>HNF4α<br>HNF4γ<br>LRH-1<br>LXRα | MR<br>NOR1<br>NURR1<br>PPARα<br>PPARγ2<br>PR<br>RARβ<br>RORβ<br>RXRα<br>RXRγ | TR2<br>TRβ<br>VDR<br><br><br><br><br><br><br>n=23 | CAR*<br>ERRβ*<br>ERRγ<br>FXR<br>NGFIB3<br>PNR<br>PXR*<br>RORγ<br>SF-1*<br>SHP | TLX<br><br><br><br><br><br><br><br><br>n=11 |

\* No detectable expression ($C_t > 34$).
\*\* Scale of y-axis.

The expression profiles are grouped into high (~ 1), intermediate (0.01 ~ 0.2), and Low (than 0.01) where the y-axis represents *normalized value* as described in methods. The asterisk (\*) represents receptors in absence.

Table 2.3 Demographic Features of Patient Samples

| Feature | Cohort (n=30) |
|---|---|
| Age -- yr | |
|     Median | 67 |
|     Range | 46-82 |
|     Mean | 65.7 |
| Gender | |
|     Female | 14 |
|     Male | 16 |
| Race | |
|     White | 27 |
|     Black | 2 |
|     Asian | 1 |
| TNM Stage | |
|     I | 16 |
|     II | 5 |
|     III | 5 |
|     IV | 4 |
| Tumor type | |
|     ADK | 23 |
|     SCC | 7 |
| Survival | |
|     Dead | 12 |
|         Female | 5 |
|         Male | 7 |
|     Alive | 18 |
|         Female | 9 |
|         Male | 9 |
| Tobacco History | |
|     Never | 5 |
|     Ex-smoker | 23 |
|     Unknown | 2 |

Abbreviations: ADK, adenocarcinoma; SCC, squamous cell carcinoma; TNM, tumor size; N,node involvement; M, metastasis status
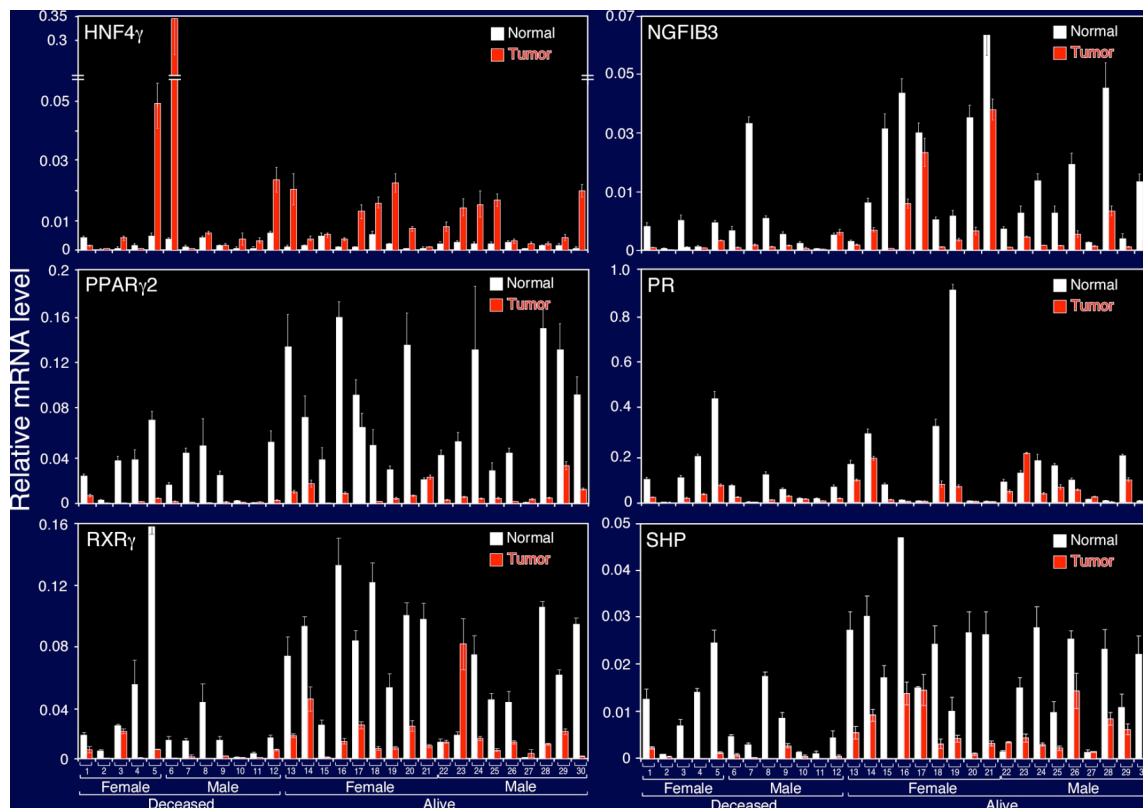
Figure 2.6 NRs Showing Dramatic Differences between Normal and Corresponding Tumor Tissues. The pair-matched samples displayed the normal tissues (white bar) and the corresponding tumor tissues (red bar). The y-axis represents the *normalized value* as described in chapter 7.
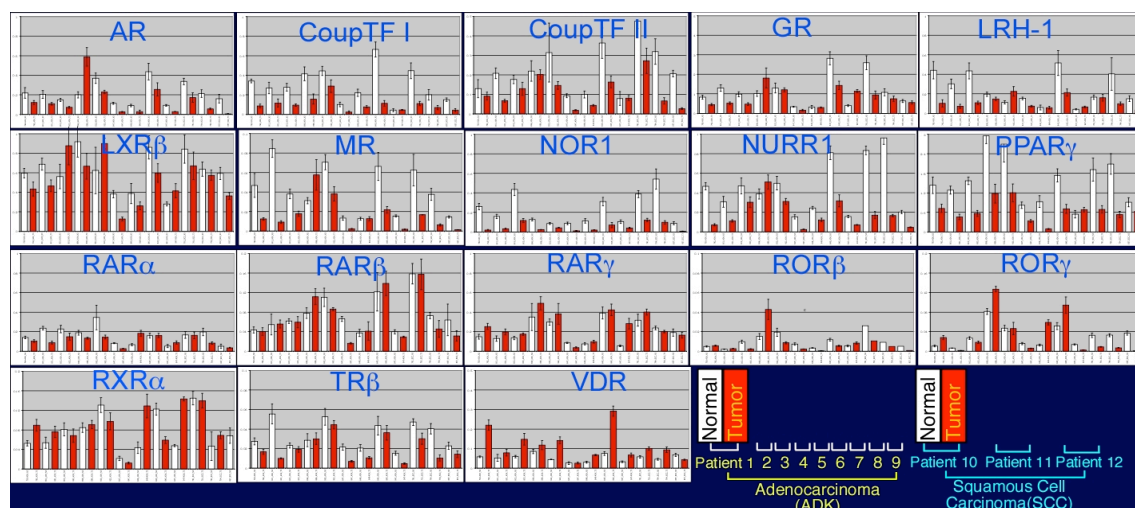
Figure 2.7 NRs showing Differential Pattern of Expression between Normal and
Corresponding Tumor in a Patient Specific Manner.
The pair-matched samples displayed the normal tissues (white bar) and the corresponding
tumor tissues (red bar). The y-axis represents the *normalized value* as described in chapter 7.
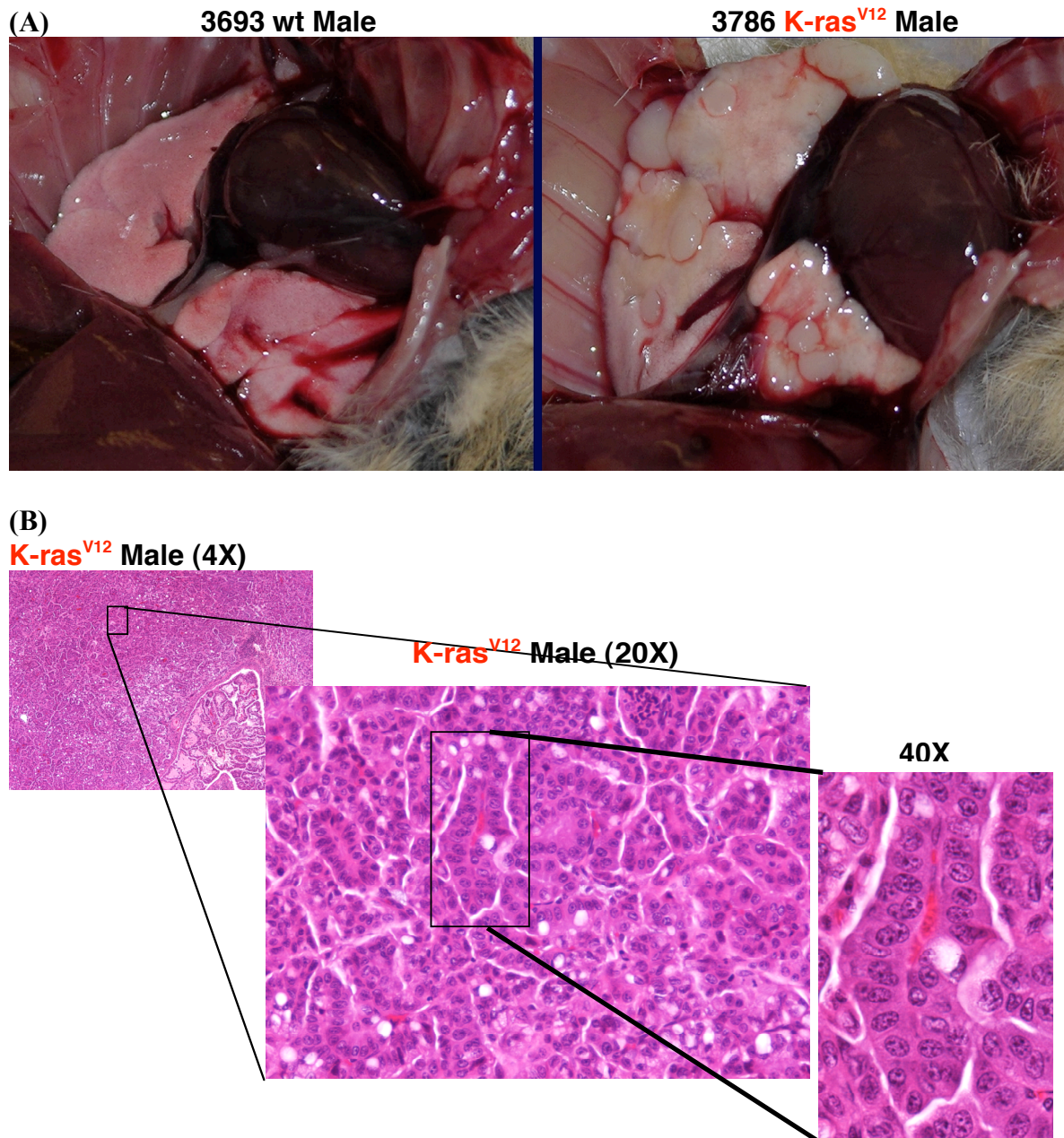
Figure 2.8 Histopathology of Mouse Lung Tumors.
(A) Pictures represent lung organ in wildtype (top-left) and *K-ras*$^{V12}$ mutant mouse (top-right). (B) H&E staining results were represented from *K-ras*$^{V12}$ lung tissues in different magnification.
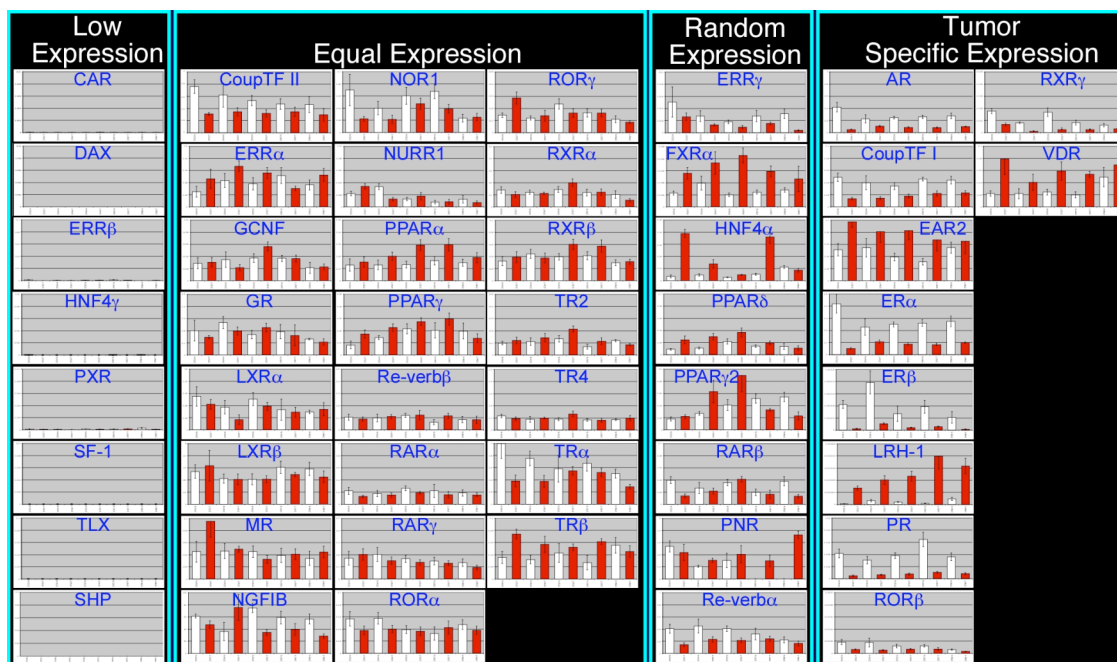
Figure 2.9 Expression Profile of NR Superfamily in Oncogenic *K-ras* Mouse Model. Expression profiling of 49 mouse nuclear receptors in normal (white) and pair-matched tumor (red) tissues from 5 male mice was classified into 4 major groups, low expression, equal expression, random expression, and tumor specific expression.

# CHAPTER THREE
## Results

### CLASSIFICATION AND PROGNOSIS

## 3.1 Introduction

Lung cancer is the leading cause of cancer related death in the United States. Even with well-known causative agents (e.g. smoke and smoke-related carcinogens), treatment options for the disease remain limited, and mostly depend on radiotherapy and surgical resection. Additionally, there has recently been successful targeted therapy in a minor percentage of cases using Iressa and Tarceva for targeting mutant EGFR. Otherwise, the last option is using combination of generic cytotoxic drugs (Bild et al., 2006; Potti et al., 2006; Thomas and Grandis, 2004). Although treatment with receptor tyrosine kinase inhibitors resulted in dramatic regression of advanced lung tumors in patients harboring mutant EGFR, patient responses vary significantly among affected populations. EGFR mutations are found at the highest frequencies in Asian non-smoking women, suggesting that further epidemiological factors are potentially involved in differential responses to the drug. Moreover, a second site mutation in EGFR results in refractory tumors against the drugs. Thus, other potential targets definitely need to be developed for therapeutic intervention. Also, the predictive value of current molecular markers has been limited. Using microarray approaches, a principal component set an, or so called 'metagene', was recently identified from a cohort training set and validated with a test set. The metagene set was shown strongly to predict lung tumor recurrence rate in the test set (Bild et al., 2006; Potti et al., 2006). In

this chapter, I will describe further analysis of the primary dataset of multiple NR profiles from various systems using bioinformatics and biostatistics tools.

**3.2 Results**

3.2.1 Diagnostic Potential of NR profiles in a Panel of Lung Cells

Most lung cancer diagnosis is made by professional pathologists who were extensively trained in the clinic. However, although this histological approach is well-established and accurate in most cases, more unbiased and objective diagnostic methods have been sought using cutting-edge technologies such as microarrays or serum proteomics to search for molecular diagnostic markers. Clearly microarray technology is a powerful tool to monitor global changes of genetic signatures of 30,000 genes at once and serum proteomics is a useful technology to identify various protein markers at a single time point. Nonetheless, the question that still remains is how to sort or handle large numbers of genes for therapeutic target validation as well as diagnostic markers. Similarly, the reproducibility and sensitivity of the proteomic approach is still of concern. Thus, keeping these issues in mind, I assessed the diagnostic potential of the NR profile in a panel of lung cell lines and a set of patient samples, as well as their prognostic relevance in a set of patient tissues.

3.2.1.1 Can the NR Profile Differentiate Normal from Tumor Cells or Tumor Types?

To investigate if the NR expression profile could be utilized to distinguish between tumor types, I performed hierarchical cluster analysis of the NR dataset using Matrix 1.29 software, a multipurpose bioinformatics tool to analyze various types of array data and to

execute unsupervised clustering analysis. The analysis revealed that the NR profile

differentiated NHBEC and immortalized HBECs from NSCLC, and NSCLC from SCLC,

although a few SCLCs were dispersed into the NSCLC cluster (Figure 3.1). The clusters

defined by profiling the 48 NRs closely resemble those obtained using 10K genes from the

microarray studies (Figure 3.2A). Consistently, the NR dataset extracted from microarray

studies showed similar patterns in the clustering results (Figure 3.2B). Moreover, bootstrap

analysis, a statistical tool that measures the clustering accuracy based on random shuffling of

the NR order, shows statistical significance in the clustering accuracy of the NR dataset

(Figure 3.3). This analysis suggests that there is strong predictable power of the NR profile

for tumor typing. The next question was to see if there is any difference in the NR expression

pattern between cell types. Since clear differences were observed in NR expression in normal

vs. tumor cells, I defined what NR subsets are altered, and how NRs are involved in the

pathogenic process. Thus, the NR expression values of individual cell lines normalized by

the average of 5 normal HBEC (NHBEC) values were formatted into a color-coded matrix

(Figure 3.4). Although it is less dramatic than the cancer cells, the immortalized HBEC cell

lines showed altered expression for certain receptors with a decreased expression when

compared to the NHBECs. However, interestingly, the cancer cells show clearly altered

expression of receptor levels. Introduction of either cellular or viral oncogenes did not seem

to induce changes in NR expression to the extent seen in lung cancer cell lines, suggesting

that other events are required for the deregulation of NR expression prerequisite for disease

progression (Figure 3.4). Further cluster analysis revealed two major groups of receptors that

are highly relevant to reproduction (Cluster I) and nutrient metabolism (Cluster II), which is consistent with our previous observation (Bookout et al., 2006).

3.2.1.2 Identification of a Subset of NRs with Diagnostic Power

To evaluate if the expression pattern of a subgroup of NRs could predict major lung cancer histological types within the various cell lines, we performed exclusion analysis and found subsets of receptors specifically expressed in different cancer types. For example, peroxisome proliferator-activated receptor gamma (PPARγ) expression was exclusive to NSCLC, consistent with previous reports (Grommes et al., 2004; Kim et al., 2005b). Notably, mineralocorcoticoid receptor (MR) also shows a NSCLC pattern of expression. Also included are expression of estrogen-related receptor β (ERRβ); estrogen-related receptor γ (ERRγ), nerve growth factor induced gene B3 (NGFIB3), progesterone receptor; retinoic acid-related orphan receptor α (RORα), which cell showed significantly higher expression in SCLC compared to NSCLC. Moreover, receptors primarily expressed in SCLC include retinoid X receptor γ (RXRγ), steroidogenic factor 1 (SF-1), and short heterodimer partner (SHP). In contrast, retinoic acid receptor γ (RARγ), thyroid hormone receptor β (TRβ), and vitamin D receptor (VDR) show moderately or dramatically reduced expression in SCLC compared to NSCLC or HBECs (Figure 3.5). Collectively, these data define predictable groups of NRs that are potentially involved in lung tumor type-specific pathogenesis. To push the NR expression signature to the next level of predictive value, we analyzed differences in expression in squamous cell carcinoma (SCC) vs. adenocarcinoma (ADK) and found that a group of receptors, e.g., estrogen receptor α (ERα); hepatocyte nuclear factor 4 γ

(HNF4γ); liver receptor homolog-1 (LRH-1); MR; PPARγ2; PPARγ; retinoic acid-related orphan receptor γ (RORγ) showed significantly increased expression in ADK vs. SCC (Figure 3.6).

Together, these data strongly support that the NR superfamily is one of the principal component marker sets that predicts lung cancer type (normal vs. NSCLC vs. SCLC) and thus may be useful in lung cancer diagnosis.

## 3.2.2 Diagnostic Potential of NR profile in a Lung Cancer Patient Panel

Since the NR profile showed diagnostic power in the lung cell panel, we next sought to identify NRs that could be clinically utilized for diagnostic purposes. Likewise, the unsupervised clustering analysis was performed with tissue samples from patients to see if any unique pattern of clustering could be generated based on tumor types (ADK vs. SCC) or demographic features. Given the same genetic background between normal tissue and the corresponding tumor tissue, one expectation might have predicted there would be 30 pairs of individual clusters. Instead, however, this analysis revealed a nice separation of the normal tissue group from the tumor tissue group (Figure 3.7). This implies that common genetic alterations were involved in the lung tumorigenic process. A further interpretation is that stem cells with preexisting alterations (e.g. cancer stem cells) develop tumors, an idea that is supported by recent identification of bronchial alveolar stem cells (BASCs) from lung a cancer mouse model (Kim et al., 2005). Furthermore, to identify genetic changes on a whole genome scale, microarray experiments using Ilumina arrays were executed for 18 patient tissue samples, which consisted of 7 normal and 11 tumor tissues. The cluster analysis

revealed 794 genes (157 genes low but 637 genes high in normal tissues) that show more than a 2-fold difference between normal and tumor groups. Consistently, tumors were in a separate group from normal tissues (Figure 3.9). To investigate if a subset of NRs rather than single NR was responsible for the difference in expression signatures between tumor and normal groups, Pearson correlation coefficients for each receptor were calculated for all receptor combinations in both normal and tumor tissues. When the resulting axis of the unsupervised cluster of NRs from tumor samples was applied to the normal group, it revealed multiple sets of NRs that lost the correlation network in tumors, while half of the NRs were still in the same cluster (Figure 3.8). Thus, it is of interest to investigate if this NR set is relevant for stem cell maintenance or tumor differentiation. Next, I further analyzed differences in expression in squamous cell carcinoma (SCC) vs. adenocarcinoma (ADK) and found that a group of receptors, e.g., AR, COUP-TFβ, COUP-TFγ, ERα, ERRγ, FXR, MR, NGFIB3, NOR1, PPARγ2, Rev-erbβ, RORβ, RXRβ, RXRγ, SHP, TRα, show significantly increased expression in ADK vs. SCC (Figure 3.10). Interestingly, analyses of lung cell lines and patient tissues revealed all receptors that significantly differentiating ADK vs. SCC were highly expressed in ADK (Figure 3.6 and Figure 3.10). Given that expression of many NRs begins with differentiation of tumors or tumor cells in colon cancer, it would be interesting to see if SCC is less differentiated compared to ADK (Gupta et al., 2003; Horkko et al., 2006; Sarraf et al., 1998). In the classification of NRs in patient samples as previously displayed in section 2.2.2, HNF4γ, which is differentially expressed in tumor vs. pair-matched normal tissue, could become a common molecular target to treat most lung cancers using gene therapy as well as a diagnostic marker along with other five receptors. On the other hand, the

NR group showing individual variations is believed to be the best fit for 'tailored drug-design', which differentiates treatment options on the basis of the NR profile. Until clinically approved, using the same approach, the NR profile in mouse model shown in Figure 2.9 is useful for the same purpose, as well as for its potential evaluation of chemoprevention.

3.2.3 Prognostic Relevance of NR Expression in Patient Samples

To identify NRs with prognostic power, we further interrogated the NR profile based on demographic features of the patients as depicted in Table 2.2. I used two independent biostatistical analyses, log rank test plotted as a Kaplan-Meier plot and a multivariate Cox regression model, which consider various factors that potentially influence survival based on the patient survival information. The NR profile was interrogated to identify a subset of NRs, if any, highly relevant to survival (Figure 3.11 and Table 2.2). The non-parametric analysis revealed multiple NRs, PR (p=0.0283), RORγ (p=0.0241) and SHP (p=0.0375), illustrated as Kaplan-Meier plots of two patient groups (one higher and one lower than the median NR value). The Cox regression model which included multiple covariates, e.g., age, gender, smoking history, histopathologic stage, recurrence, tumor type, were performed next for the entire 48 NR set. This semi-parametric analysis revealed that AR, PR, and RARβ are associated with a low hazard ratio (less than 1) and therefore a good prognosis for survival when highly expressed in the corresponding tumor, whereas PPARδ displayed a high hazard ratio (greater than 3), meaning a positive correlation to tumor progression (Table 3.1). Interestingly, none of the clinical factors appears to be significantly relevant to survival in the multivariate analysis. In addition to the log-rank test and Cox regression model focusing on

individual NR expression for patient survival, principal component analysis was used to

identify NR subsets (i.e., metagene) that are most reflective of the entire dataset and their

pathophysiological meaning still remains to be investigated (Figure 3.12). Note that both

principal component sets (1 and 2) show no correlation to patient survival (Figure 3.12B). On

the other hand, PC set 3 containing SF-1 and TLX, did correlated very well with survival as

evidenced by Kaplan-Meier plots (Figure 3.12C). Importantly, neither TLX or SF-1 alone

had this predictive power, emphasizing the utility of this type of analysis. Recently, Chen et

al., published prognostic evaluation of a five-gene signature in clinical NSCLC samples.

From microarray data analysis of 125 NSCLC tissue samples, risk scores were calculated for

each of the five genes. Patient risk scores are a summation of risk scores from the individual

five genes, which are a product of expression level and hazard ratio from univariate Cox

regression model (Chen et al., 2007). We applied the same biostatistical analysis to the

expression profile of the nuclear receptor superfamily. Strikingly, the first 10 nuclear

receptors with high risk scores (or Z scores) revealed highly significant prognostic potential

(Figure 3.13A), whereas the intermediate 10 NRs and the last 10 NRs showed no

discriminating power for patient survival. Moreover, five NRs were identified from further

multivariate Cox regression of the first 10 genes and shown to provide a prognostic power

similar to all 10 NRs (Figure 3.13B).

3.3 Discussion

The NR atlas has been assessed for potential application to providing diagnostic and

prognostic information in various systems including lung cell lines and patient tissue

samples. First, in the cluster analysis of histopathologically classified cell lines and patient tissue samples, the NR atlas revealed strong predictable power for distinguishing different tumor types and revealing the degree of carcerogenesis progression from normal tissues to tumors (Figure 3.1 and Figure 3.7). Interestingly, receptors with common regulatory rules appeared to be grouped together in the cluster analysis (Figure 3.1 and Figure 3.4). These data suggest that homeostatic dysregulation (e.g., in metabolic pathways) caused by malfunction of multiple NRs may result in chronic diseases like cancer. Furthermore, a clear segregation of normal tissues from tumor tissues in the cluster analysis supports the idea that a generation of lung cancer stem cells may result in part due to the failure of maintaining a physiological balance by a subset of NRs (Figure 3.7).

Although NR profiling of lung cell lines and show patient tissues were able to classify tumor types, and normal from tumor tissue, the datasets from the lung cell lines were not similar to the patient tissue profiles. In other words, the expression profile from tissues must be performed independently for clinical application. Nevertheless, the cell line data does have enormous utility. Expression of certain NRs in the cell line clusters reflects the pathological progression of normal to immortalized, and were on to lung cancer (Figure 3.4). Multiple NRs (green colored in cluster II) involved in gain of immortality (after introducing CDK4 and hTERt). This subset tends to be similarly maintained even after oncogenic changes (e.g. human papilloma virus E6/E7, oncogenic K-ras, p53 knock-down), which resulted in moderate changes for some NRs in cluster I. Interestingly, BEAS2B cells harboring SV40T antigen displayed relatively advanced changes for a few NRs. However, the balance of NR expression was dramatically different in lung cancers (Figure 3.4).

Furthermore, amongst these NRs most predictive of tumor type – ERRβ, ERRγ, NGFIB3, PR, RORα, RXRγ, SF-1, and SHP for SCLC; and MR, PPARγ for NSCLC; and RARγ, TRβ and VDR for SCLC, it is of particular interest that both SHP and RXRγ are known to regulate the functions of other receptors by direct binding. SHP is an atypical nuclear receptor with no DNA binding domain, which negatively regulates the activity of other NRs (Lee et al., 1998; Lu et al., 2000; Seol et al., 1998). In contrast, RXRγ, one of three RXR subtypes that functions as an obligatory binding partner, is believed to positively regulate the activity of other NRs (Shulman et al., 2004; Shulman et al., 2005). It will be interesting to see if physiological malfunction of these two receptors has pathological relevance to SCLC development. The decreased expression of VDR in SCLC implicates a potential role to prevent SCLC development. The expression of PPARγ in NSCLC suggests that pathogenesis of this particular tumor type is relevant to fatty acid metabolism as well as to the anti-inflammatory role of the PPARγ. Also, the combined treatment of a PPARγ agonist, pioglitazone, with an antagonist of MR, which is known to activate the inflammatory pathway by aldosterone binding, may be a worthwhile scheme to treat lung adenocarcinoma.

Further analysis of NSCLC revealed groups of receptors differentially expressed between adenocarcinoma (ADK) and squamous cell carcinoma (SCC). Four of these receptors ERα, HNF4γ, LRH-1, and RORγ, may also be used for predicting the ADK phenotype. Interestingly, SCC specific-expression of NRs was not observed in either the panel of lung cell lines or patient samples, implicating that loss of functional activity of receptors might cause the development of the SCC phenotype. In addition, the ADK-specific NRs were not shared between the lung cell panel and tissue panel, suggesting again that the

lung cell lines, although still useful for preclinical evaluation, are not representative of patient tissues, which may be due in part to the heterogeneity and microenvironment of primary tumor tissues.

In comparing the cluster analysis between tumor and normal tissues, three distinct subsets of NRs found in normal tissues were lost in the tumor samples. These subsets included NRs VDR, PPARδ, PPARδ2, PPARγ2, LXRα, RARα, RXRα and γ that play roles in nutrient metabolism and NURR1, NOR1, and NGFIB3 that are early-response members for numerous physiologic processes. This suggests that crosstalk within subgroups of NRs may play important roles in maintaining the normal phenotype of lung tissue. It will be crucial to investigate if similar changes or common critical changes of NR subsets occur in all adenocarcinomas from various tissues. This may provide insight into the identification of an unified pathological mechanism for cancer incidence, which could then be utilized for treatment and/or chemoprevention. On the other hand, the cohort of NRs that remain clustered in both normal and tumor tissues may be involved in maintaining cell survival. Thus, it may be interesting to investigate what physiological changes may occur if this group is disrupted. The pathologic process may be involved in loss of control not only by a single receptor but by a group of NRs that function as a network and play an important role in sustaining physiological homeostasis of lung tissues.

Along with the pathologic relevance of NRs, further prognostic evaluation using biostatistics analysis (e.g., multivariate regression model) revealed multiple NRs of interest, including AR, PR, PPARδ, and RARβ, have significant relevance to patient survival. It is important to note that colleagues at Duke revealed no significant correlation of AR to lung

cancer prognosis when analyzed in a cohort of tumor tissues alone (data not shown). However, it turns out that AR expression is significantly correlated to patient survival when analyzed in pair-matched tissue samples that could exclude individual variation due to genetic background. Thus, our data represent a unique set that may have additional predictive power to other datasets from tumor-only samples. RARβ expression, consistent with previous studies for its positive prognostic value, displayed a low hazard ratio in tumor expression compared to other covariates. Likewise, PR expression, which is an already well-known prognostic marker and even therapeutic target in breast cancer, also revealed a significantly low hazard ratio in tumors. Strikingly, PPARδ showed a high hazard ratio (meaning a high risk factor for lung cancer). It has been proposed that PPARδ is involved in tumor cell growth in cell culture systems (Fukumoto et al., 2005; Han et al., 2005; Wang et al., 2004a). Considering that PPARδ is involved in fat metabolism, it will be of interest to study if or how fatty acid metabolism is relevant to lung cancer incidence. Further identification of five NRs, including PPARδ, PR, HNF4α, VDR, and LXRα, from the risk score summation analysis, provides a novel set of NRs for prognostic purpose as well as therapeutic targets.

Overall, a thorough analysis of the NR profile revealed striking potential for various aspects of translational research including diagnosis, prognosis, and even pathologic relevance.
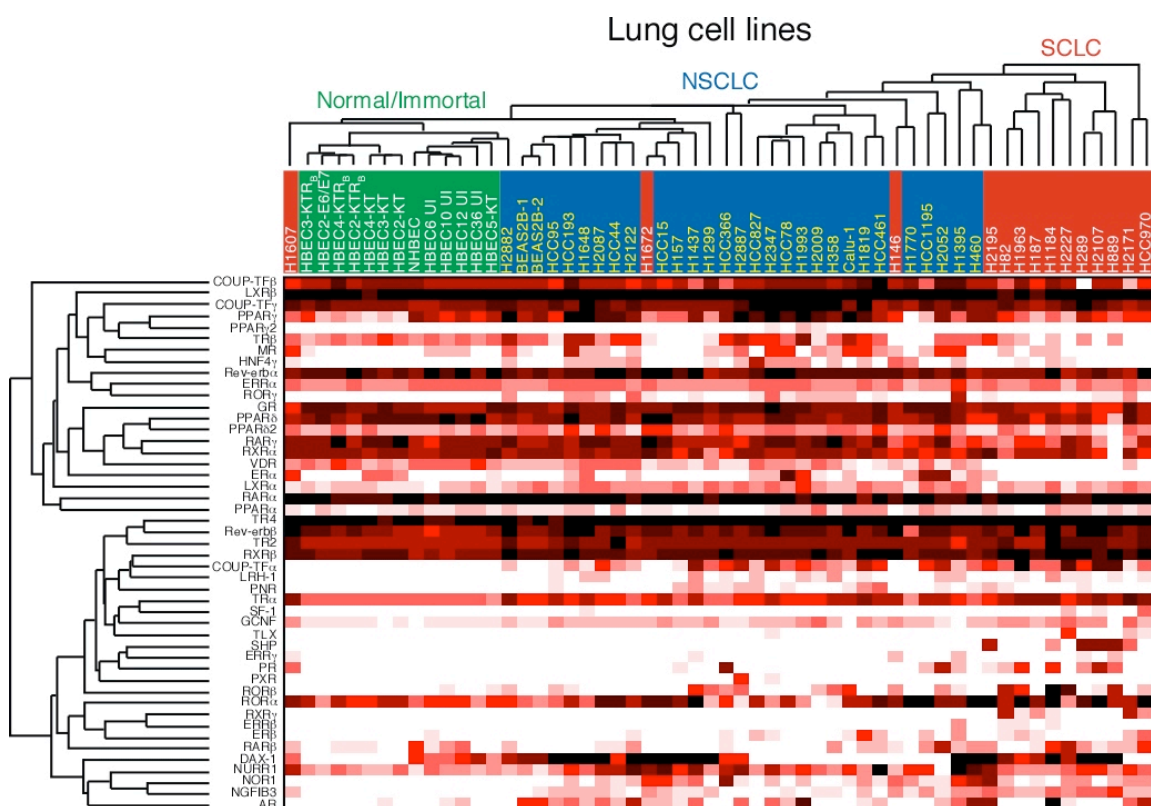
Figure 3.1 Unsupervised Cluster Analysis of Nuclear Receptor Profile.
Unsupervised clustering analysis for both 56 lung cell lines on horizontal axis and 47
NRs on vertical axis. QPCR values are clustered using Matrix 1.29 software and color-
coded according to the expression level. Green colored group represents immortalized
and normal human bronchial epithelial cells. Red colored group represents SCLCs, and
blue for NSCLCs or ADK in the bottom panel. The scaled bar is relative $\log_2$ ratio value.
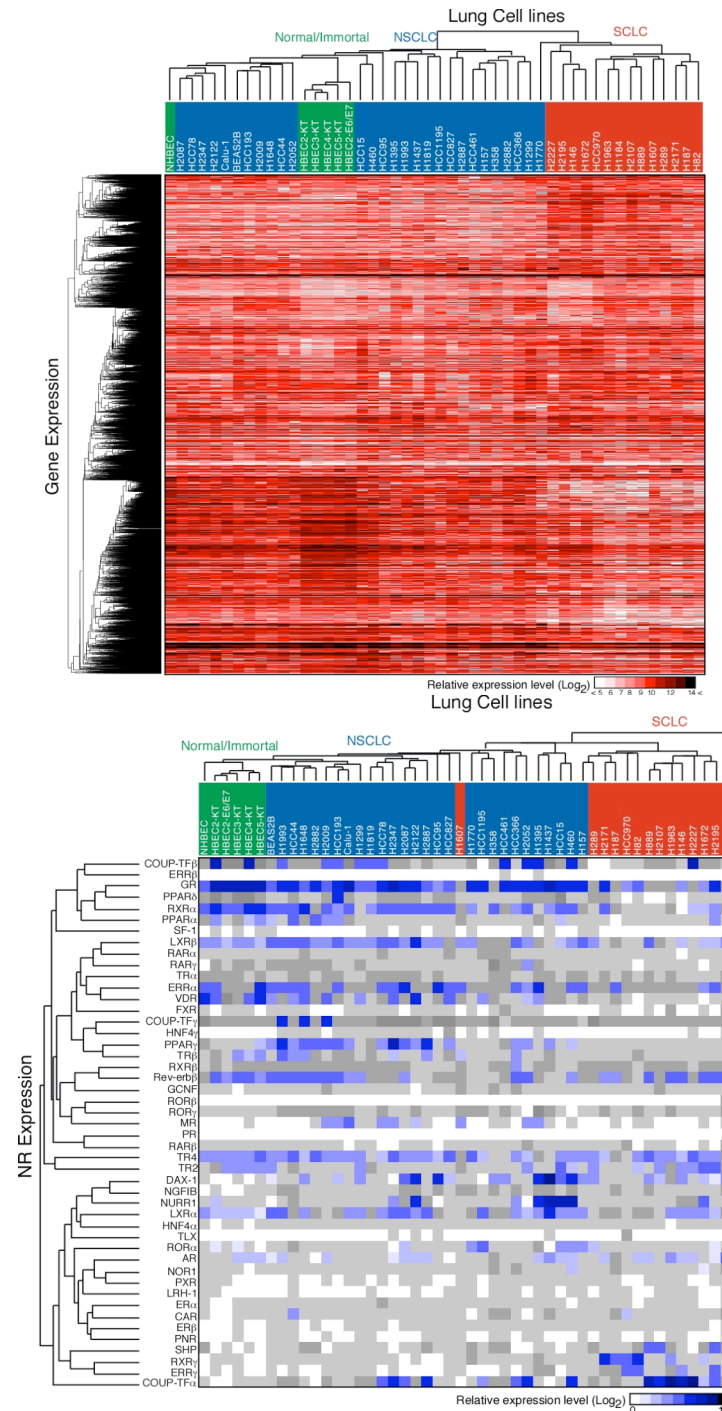
Figure 3.2 Unsupervised Cluster Analysis of 10,000 genes and the NR Superfamily from Affymetrix Microarray. (A) Hierarchical clustering of 47 NRs from the affymetrix Microarray data. The average values of multiple probes for each nuclear receptor in the microarray data were used for the clustering analysis. Note that probes specific for PPARδ2, PPARγ2, and Rev-erbα are not present on Affychip. (B) Unsupervised clustering analysis for both 48 lung cell lines on horizontal axis and genes on vertical axis. Approximately 10,000 genes filtered out of 47,000 genes on Affymetrix U133AB chip are utilized to perform the unsupervised clustering analysis both of genes and lung cell lines listed. A group of green represents NHBEC and immortalized bronchial epithelial cells, red for SCLCs, and blue for NSCLCs.
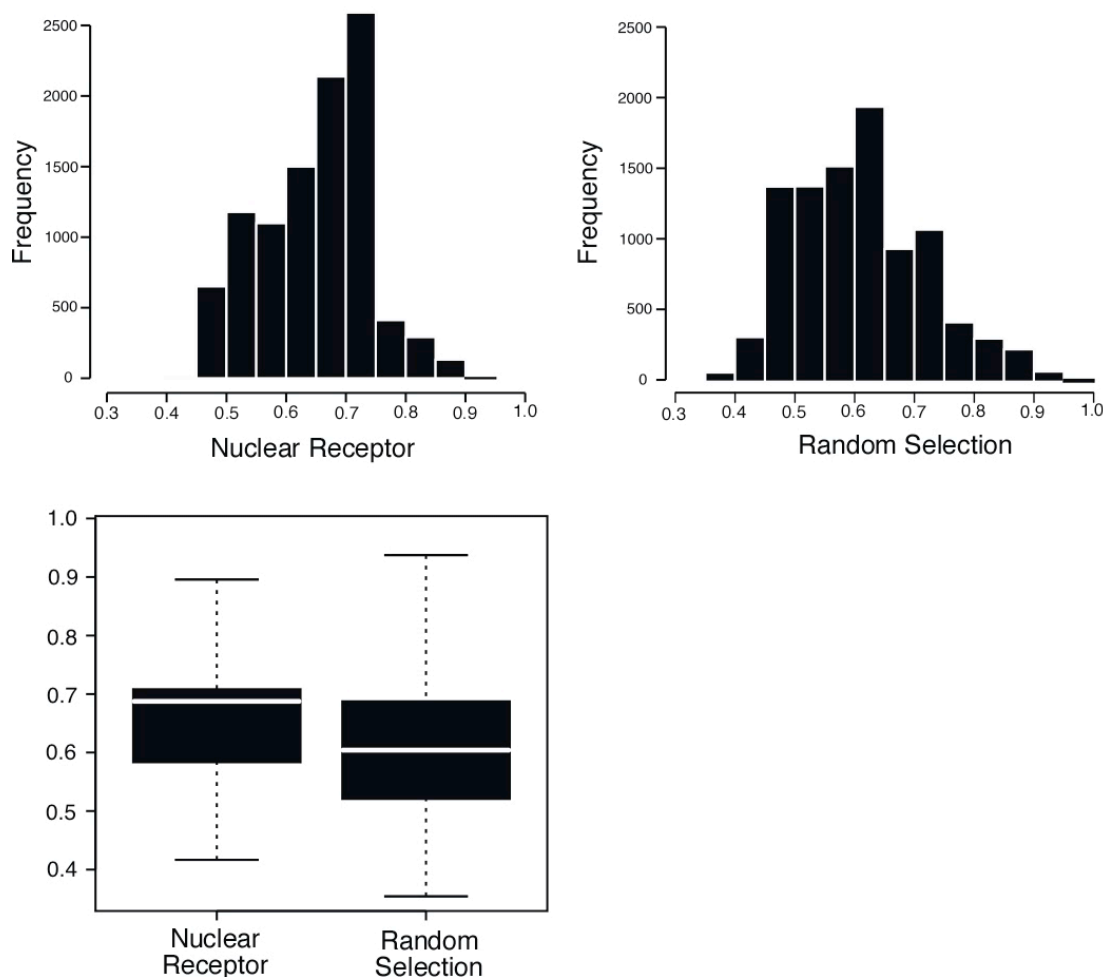
Figure 3.3 Statistical Significance of NR Superfamily for the Clustering Analysis. Bootstrap analysis was performed using a Python script was written using Python® (http://www.python.org) for bootstrap analysis, a statistical method measuring clustering accuracy. Both 123 probes set of nuclear receptors and same number of randomly selected probe sets from the microarray data were tested for the clustering accuracy. The PyCluster, a python module implemented K-means method, is used for the clustering analysis. The x-axis represents the accuracy of the clustering and the y-axis represents the frequency in the histogram plot. The bold line in box plot stands for the median value of the clustering accuracy. The statistical analysis using R package for both *u-test* and *t-test* reveals highly different between both bootstrap analyses (*P< 2.2e-16*) (http://www.r-project.org).
(Top) Histogram of nuclear receptor probe sets (NR) and randomly subsampled
   probe sets (RS).
(Bottom) Box plot for median of clustering accuracy.
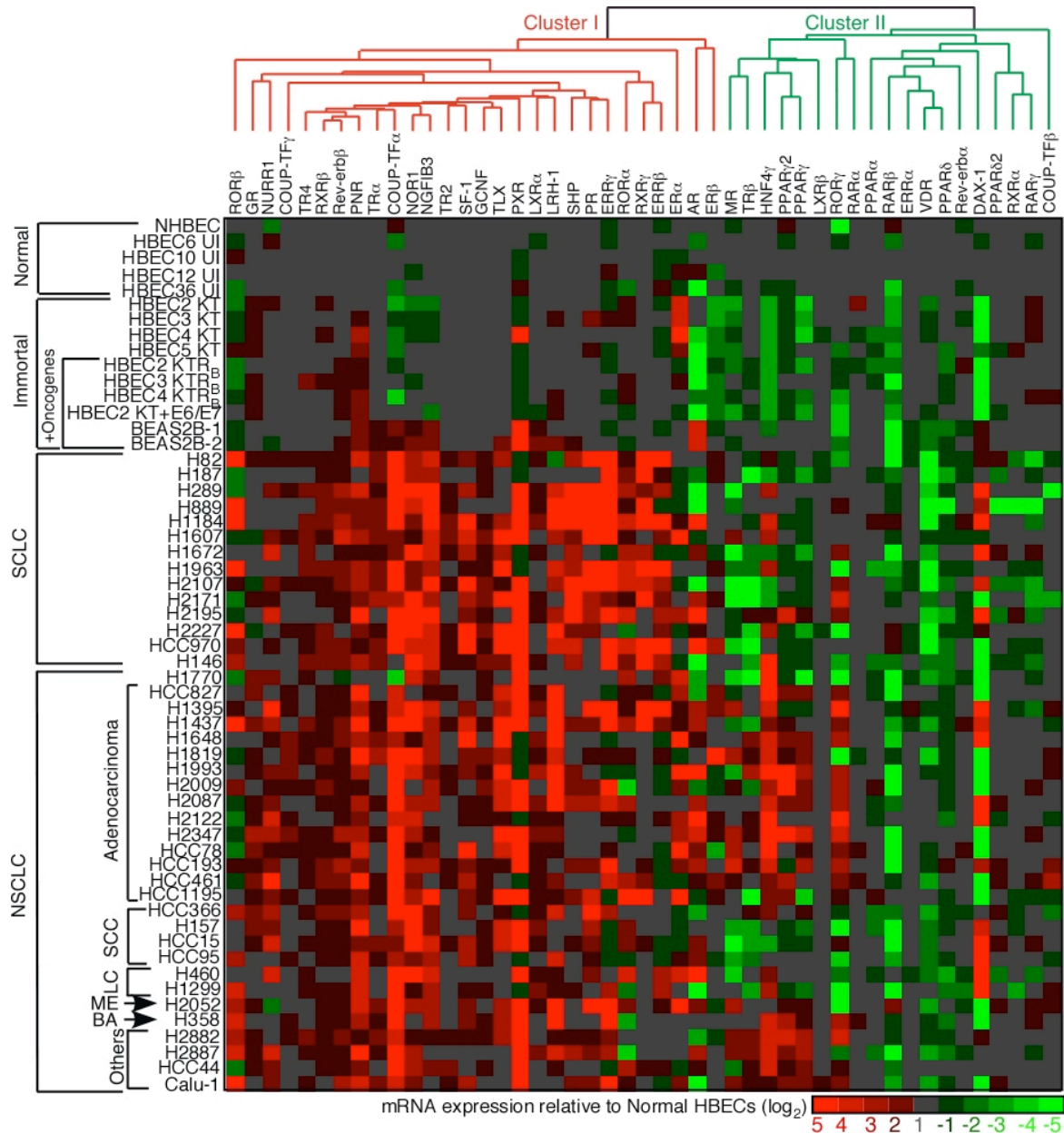
Figure 3.4 Pathologic Comparison of Nuclear Receptors from Normal to Immortalized, to Lung Cancer Cells. Color-coded signature for receptor expressions of the lung cancer or the immortalized cell lines normalized by the average value of each receptor in five normal epithelial cells. The horizontal axis shows the unsupervised clustering of 47 NRs and the vertical axis displays 56 lung cell lines on the pathological classification. NHBEC, normal human bronchial epithelial cell; HBEC UI, HBEC uninfected cells; SCLC, small cell lung carcinoma; SCC, squamous cell carcinoma; NSCLC, non-small cell lung carcinoma; ME, mesothelioma; BA, bronchioalveolar. Note that CAR, FXR, HNF4α are outliers due to low expression. The scaled bar represents $\log_2$ ratio of the normalized values by normal HBECs.
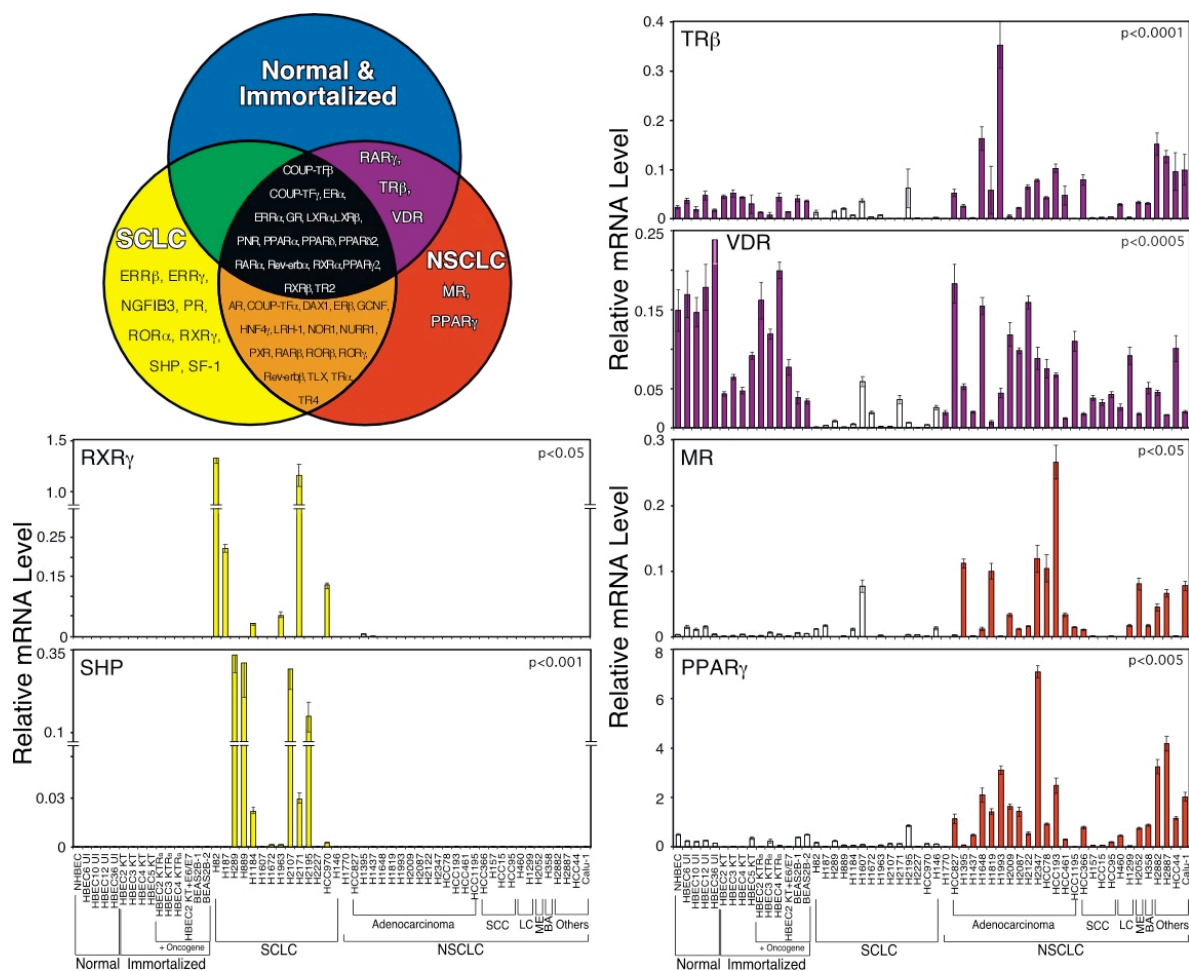
Figure 3.5 Identification of a Subset of NRs that differentiate SCLC from NSCLC. A subset of receptors is grouped in the *Venn diagram* to distinguish expression pattern on the statistical *U-Test*. The bottom panels display individual receptors identified from the above classification. The NRs are represented for SCLC with yellow, both normal/immortalized cells and NSCLCs with purple, and NSCLC with red. Note that CAR, FXR, HNF4α are outliers due to low expression.
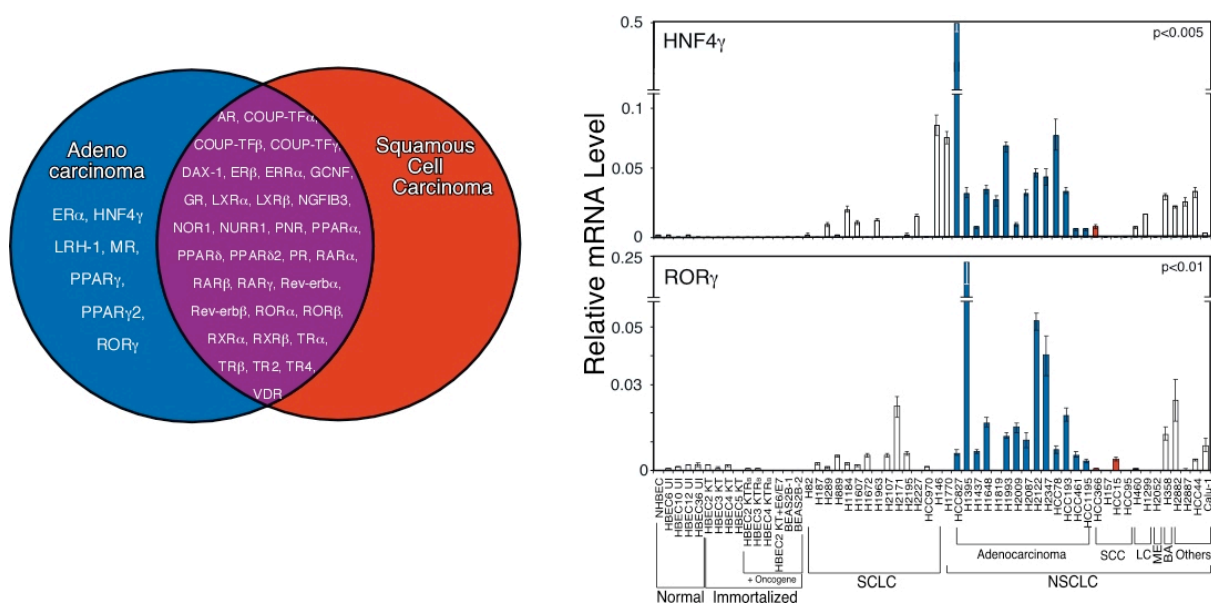
Figure 3.6 NRs Differentiating Adenocarcinoma (ADK) from Squamous Cell Carcinoma (SCC). A *Venn diagram* shows distribution of NR expression between adenocarcinoma and squamous cell carcinoma. The bar graphs display mRNA expression of the individual receptor showing distinct pattern of expression according to the tumor types. A group of bar graphs with blue or red represents ADK or SCC, respectively. For statistical significance, statistical *U-test* was performed between the two groups. Note that CAR, ERRβ, ERRγ, FXR, HNF4α, PXR, RXRγ, SF-1, SHP and TLX are excluded due to low expression in these two groups of cell lines.
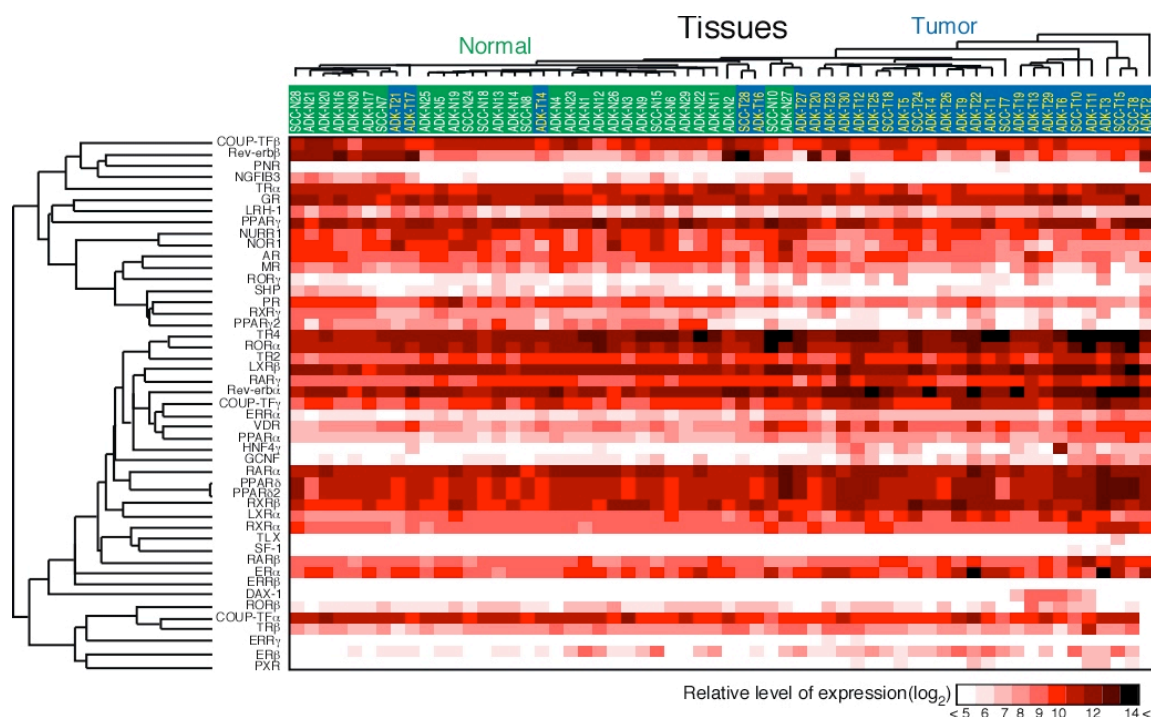
Figure 3.7 Unsupervised Cluster Analysis of Nuclear Receptor Profile. Unsupervised cluster analysis of patient tissues (horizontal axis) and 47 NRs (vertical axis). QPCR values were clustered using Matrix 1.29 software and color-coded according to expression level. Green represents immortalized and normal human bronchial epithelial cells. Red represents SCLCs, and blue for NSCLCs or ADK. The scale bar is relative $\log_2$ ratio value.

Figure 3.8 Comparison of Statistical Clustering Analysis between Normal and Tumor Tissues. Pearson correlation coefficient of one receptor to another in a panel of tissue samples was calculated and converted into color-codes, i.e., negative correlation in green and positive correlation in red between two receptors. The same order of NR axis from hierarchical cluster analysis in a set of tumor tissues (left) was applied to a set of normal tissues. Four receptors, CAR, ERRβ, PNR, and SF-1 were outliers due to low expression. The scaled bar represents Pearson correlation.
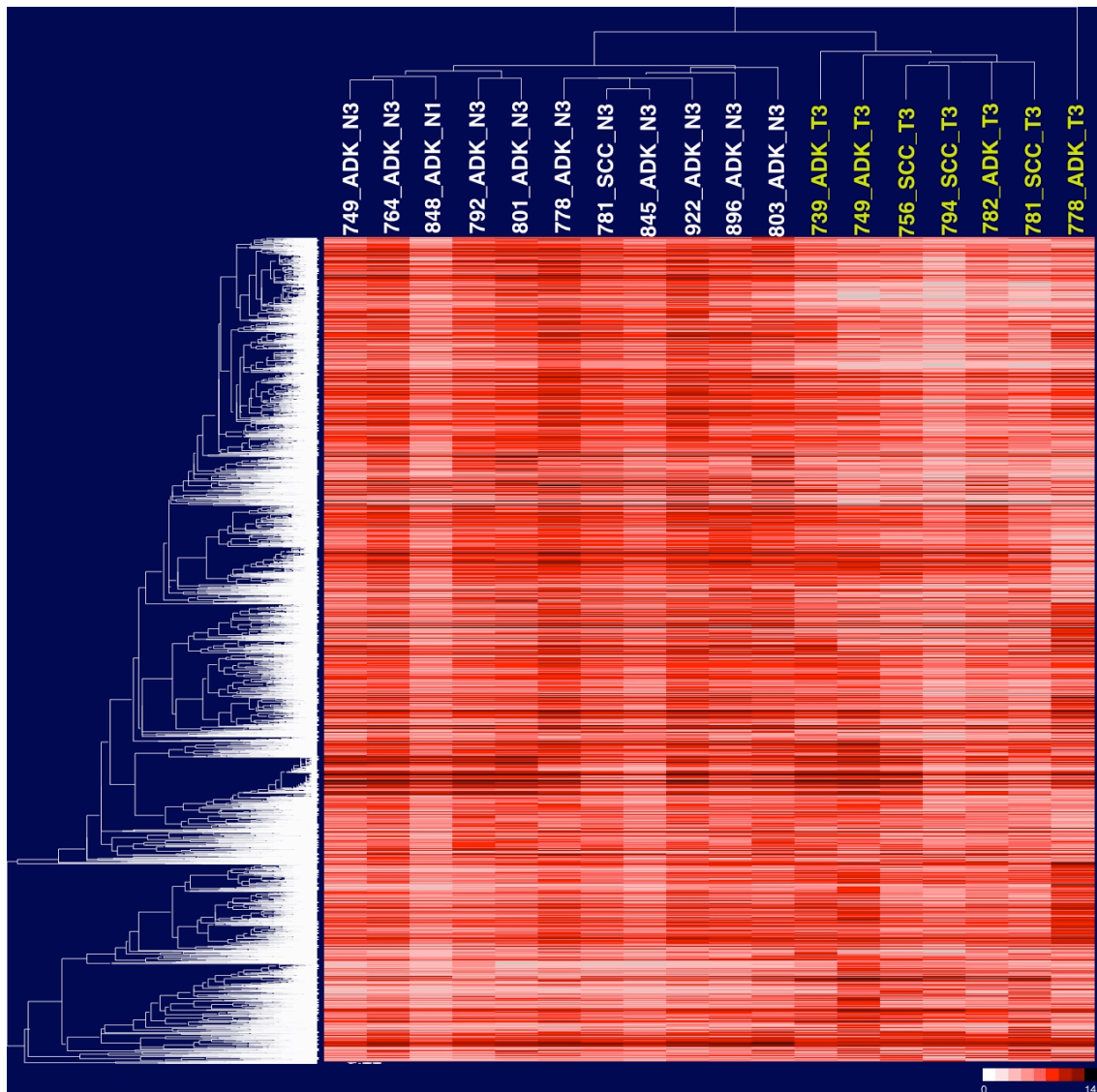
Figure 3.9 Expression Signatures of Normal and Tumor Tissues Using Ilumina Arrays. Genetic signatures in 18 lung tissue samples (11 normal and 7 tumor tissues) were surveyed using Ilumina microarrays. Further cluster analysis was performed for both samples (horizontal) and gene expression (vertical).
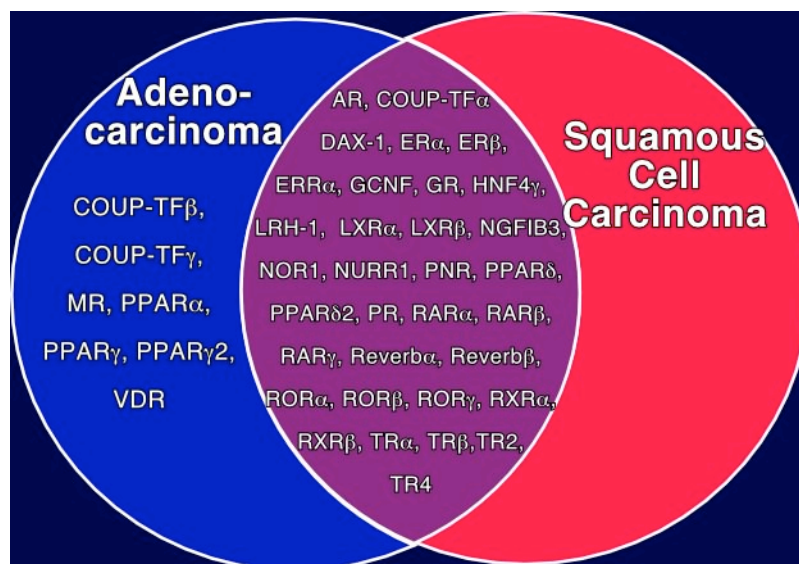
Figure 3.10 NRs Differentiating ADK and SCC in Patient Tissues.
A *Venn diagram* shows distribution of NR expression between adenocarcinoma and squamous cell carcinoma. For statistical significance, statistical *U-test* was performed between the two groups. Note that CAR, ERRβ, ERRγ, FXR, HNF4α, PXR, RXRγ, SF-1, SHP and TLX are excluded due to low expression in these tissue samples.

Table 3.1 Nuclear Receptors Associated with Survival based on a Cox Regression Model.

| Nuclear Receptors | ChiSquare | |
|---|---|---|
| AR | 0.0369 | 0.0010 |
| PPARδ | 0.0084 | 3.6990 |
| PR | 0.0314 | <0.000009 |
| RARβ | 0.0205 | 0.0570 |

For the parametric models, various clinical factors, i.e., age, gender, smoking history, cancer stage, recurrence status, histology together with NR expression level, are included as multiple covariates. NRs not significantly accounted:
CAR, COUP-TFα, COUP-TFβ, COUP-TFγ, DAX-1, ERα, ERβ, ERRα, ERRβ, ERRγ, FXR, GCNF, GR, HNF4α, HNF4γ, LRH-1, LXRα, LXRβ, MR, NGFIB3, NOR1, NURR1, PNR, PPARα, PPARδ2, PPARγ, PPARγ2, PXR, RARα, RARγ, Rev-erbα, Rev-erbβ, RORα, RORβ, RORγ, RXRα, RXRβ, RXRγ, SF-1, SHP, TLX, TR2, TR4, TRα, TRβ, VDR
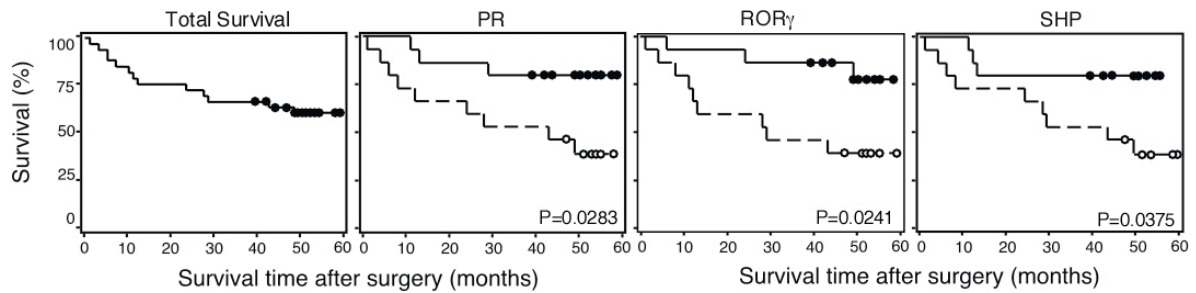
Figure 3.11 Kaplan-Meier Plot of Patients and Nuclear Receptors Associated with the Survival.
Survival plots of patients were displayed for PR, RORγ and SHP that were identified from long-rank test. Based on expression level, patients are median-splitted into two groups, high NR expressers and low NR expressers, followed by Kaplan-Meier plots for each receptor. In this analysis, the ratio of tumor value to corresponding normal value was used for each receptor. Circles in each group represent censored patients when the data were analyzed.
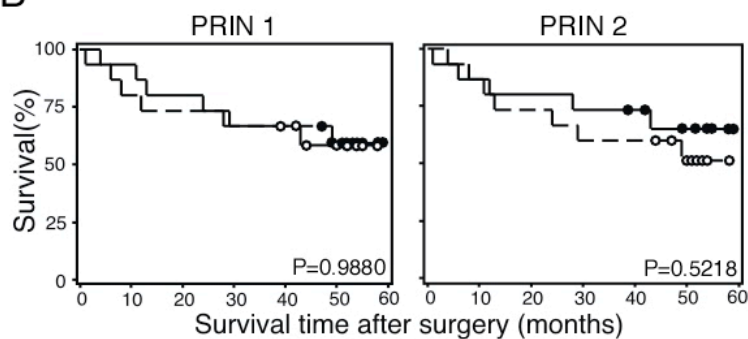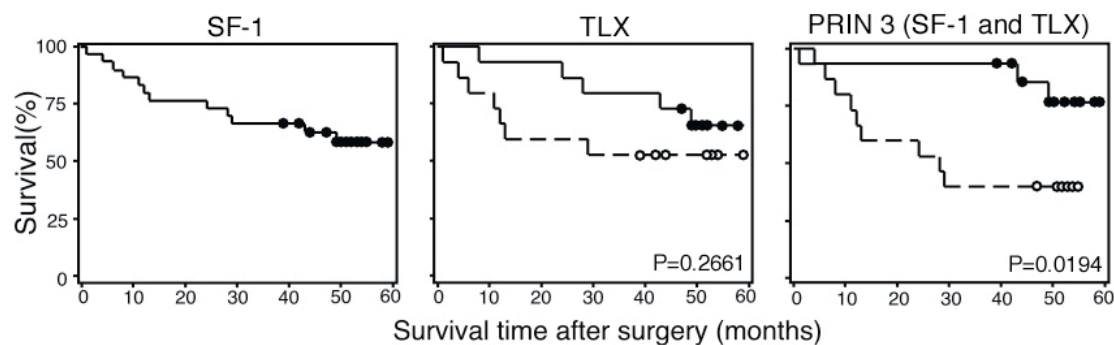
Figure 3.12. Metagene Analysis for Nuclear Receptor Superfamily.
The metagene analysis was executed using the ratio of tumor value to corresponding normal value for each receptor. (A)Three representative principal component sets were displayed in table. (B) Log-rank test of two metagenes for patient survival. (C) Kaplan-Meier plot for metagene 3 and individual receptor components.

A



B



Multivariate Cox regression model

| Nuclear Receptors** | ChiSquare | Hazard Ratio |
|---|---|---|
| PPARδ | 0.12 | 1.77 |
| PR | 0.015 | 0.000086 |
| HNF4α | 0.012 | 1.1 |
| VDR | 0.023 | 0.38 |
| LXRα | 0.019 | 1.6 |
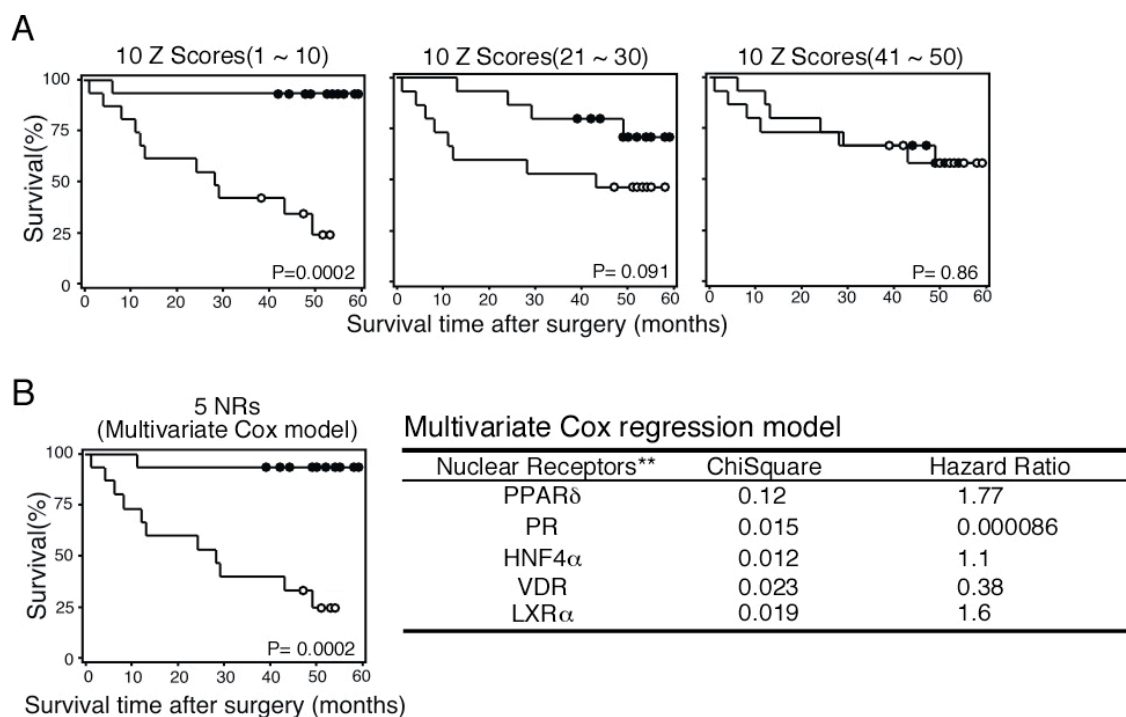
Figure 3.13 Estimates of Patient Survival to the Ten and Five NR Signatures.
Using the ratio of tumor value to corresponding normal value for each receptor, Z
scores (or risk scores) of individual patients were calculated and the survival was
represented on Kaplan-Meier plot by median-split. (A) Kaplan-Meier plots with ten NR
signatures of highest 10 Z scores (left), intermediate 10 Z scores (middle), and lowest
10 Z scores (Ramirez et al.). (B) Kaplan-Meier plots with five NR signatures. Top five
NRs from multivariate Cox regression analysis of ten NRs with the highest Z scores
(left figure in A) were used for patient survival. A table shows statistical significance
and hazard ratio of the five genes.

# CHAPTER FOUR
## Preclinical Functional Consequences of Therapeutically Relevant NRs

## 4.1 Introduction

Due to extensive biomedical research on various types of cancers, recent clinical trials have shed light on molecular targeted therapy increasing clinical potency and specificity. Some of these drugs are Gleevec® (STI-571 or imatinib), which targets receptor tyrosine kinase for gastrointestinal stromal tumor and certain kinds of chronic myeloid leukemia; Volcade® (bortwezomib, which blocks proteosome to treat multiple myeloma; and Genasense™ (oblimersen), which is used to treat non-Hodgkin's lymphoma and solid tumors by blocking Bcl-2 expression. Both Volcade® and Genasense™ induce apoptotic cell death in treated cancers. In addition, Herceptin® (Trastuzmab), a humanized monoclonal antibody against Her2/neu is used in the treatment of breast cancer. Likewise, Iressa® and Tarceva®, inhibitors of the receptor tyrosine kinase EGFR, have recently shown dramatic clinical outcomes in some lung cancer patients. However, unlike the initial promising clinical outcome with the inhibitors, the recurrence with second site mutations in the EGFR resulted in a refractory response to the same drugs. Moreover, only 10 ~ 20% of all NSCLC patients show any response to these tyrosine kinase inhibitors (TKIs). Surely, as more targets become

available, the better the clinical outcome will be. Extrapolating from the previous chapter dealing with the diagnostic and prognostic value of the entire NR atlas and NR subgroups, I next asked if receptors could be potential targets for therapeutic intervention. A subset of NRs including ERα, AR, and PPARγ were selected for further analysis because they have been extensively studied in various types of cancers such as breast, prostate, and colon cancers as discussed in section 1.3 (Allred and Kilgore, 2005; Joly-Pharaboz et al., 2000; Liu et al., 2002; Liu et al., 2003; Osipo et al., 2003; Umekita et al., 1996). Moreover, ligands for these receptors are available, and have been widely used in the clinic. Thus, various lung cancer cell lines selected for each receptor were preclinically evaluated for cell growth in response to ligand treatment and the therapeutic potential of PPARγ was further assessed *in vivo* using a mouse xenograft model.

## 4.2 Results

### 4.2.1 Androgen Receptor

Several human lung cancer cell lines expressing (e.g. H1184, H2122, H1993) or not expressing (e.g. H1299 and H2009) AR were examined for their cellular responses to the agonist dihydrotestosterone (DHT) (Figure 4.1). LnCaP, an AR positive prostate cancer cell line, was used as a positive control and showed consistent growth stimulation with DHT treatment at a physiological concentration (1 nM) as previously described (Okamoto et al., 1997; Yang et al., 2005). Two AR negative lung cancer cell lines, H1299 and H2009, showed very little or no response to the agonist treatment as expected. H1184 cells, a SCLC line that has the highest expression of AR, showed significant growth-stimulation in a range from low

physiological (0.1 nM) to physiological nanomolar concentrations (10 nM) of DHT, but growth inhibition at a supraphysiological concentration (1 μM). In addition, H2122 cells showed growth stimulation to a similar extent as H1184 cells, but seemingly had no cytotoxic response at the same pharmacological concentration. Interestingly, H1993 cells showed a growth inhibitory response to DHT treatment at all concentrations. Overall, DHT treatment displayed receptor-dependent cell growth responses, although not necessarily growth stimulatory.

4.2.2 Estrogen Receptor

Two positive expressers, H2052 and HCC78, were treated either with the agonist 17β-estradiol or with the antagonist ICI 182,780 for 3 consecutive days (Figure 4.2). Estradiol treatment at physiological concentrations (10 nM) caused a 30% increase in cell growth and 40% growth inhibition was seen with the pure anti-estrogen ICI 182,780 in H2052 cells. Note that estrogen receptor β (ERβ) is not expressed in these selected lung cancer cell lines. Breast cancer MCF-7 cells used as a positive control showed over 2-fold growth stimulation with 10 nM estradiol and approximately 50% growth inhibition with 1uM ICI 182,780 compared to the vehicle treatment. The difference in the magnitude of the growth response is possibly due to the approximately 50-fold higher expression of ERα in MCF-7 cells compared to H2052 cells (data not shown). HCC78 cells showed 30% growth inhibition at a physiological concentration of 17β-estradiol (10 nM) treatment, which is similar to the observation of AR-responsiveness in H1993 cells. However, there was no observed response to ICI 182, 780 at low concentrations. The analysis of the two steroid

receptors (e.g., AR and ERα) demonstrated that the mRNA expression reflects the functional activity shown in growth response in cell culture.

4.2.3 Peroxisome Proliferator-Activating Receptor gamma (PPARγ)

The PPARγ-specific agonist troglitazone showed significant growth inhibition in a dose dependent manner in three high expressers (e.g. H2347, H1993, Calu-1) but no significant effects in three low expressers (e.g. H1770, HCC1195, H1299). PPARγ protein levels were confirmed by western blot analysis (Figure 4.3). *Yang et al.* previously showed that activation of PPARγ induces cell cycle arrest and apoptosis in human renal carcinoma cell lines due to decreased cyclin D1 expression (Yang et al., 2005). Thus, to see if growth inhibition could be mediated by cyclin D1 inhibition, we analyzed cyclin D1 expression in two cell lines, H1770 cells with low PPARγ expression and H2347 cells with high PAPRγ expression, treated with troglitazone and found decreased cyclin D1 expression specifically in H2347 cells. These data suggest that the decreased expression of cyclin D1 with troglitazone treatment is a PPARγ dependent cell growth inhibitory effect (Figure 4.3). Next, to validate the potential therapeutic effects of a PPARγ ligand *in vivo*, mouse xenografts were established by subcutaneous injection of two million H1770 or H2347 cells into the flank of immune-compromised nude mice. A dose of 25mg/kg of pioglitazone, a PPARγ agonist commercially available to treat Type II diabetes, was intraperitoneally administered four times a week for 5 weeks, starting 5 days after tumor cells were injected as described in the methods section. The pioglitazone treatment had no effect on H1299 tumor growth, and even seemingly promoted H1770 tumor growth in two out of four mice compared to the vehicle

treatment, consistent with the observation with in vitro troglitazone treatment. However,

H2347 cells showed a clear inhibition of tumor growth with pioglitazone treatment compared

to vehicle treatment, implicating the receptor dependent therapeutic potential of this PPARγ

ligand for lung cancer treatment. Interestingly, the tumor-promoting effect of pioglitazone in

receptor negative H1770 cells suggests that therapeutic application should be performed

based on the expression of target receptor for treating lung cancer (Figure 4.4).

4.2.4    Differential Gene Expression on Receptor Expression

To further explore the molecular changes in lung cancer cell lines occurring

according to expression of the steroid receptors and PPARγ, Affymetrix U133AB array

experiments were performed in a large panel of 48 lung cell lines. Note that the culture

condition for maintaining these lung cancer cell lines includes 0.1 nM of 17β-estradiol which

is a physiologically relevant concentration (refer to growth response of MCF-7 in Figure 4.1)

and 0.005 nM of testosterone that could induce the basal activity of AR in culture. Indeed,

microarray data analyses show 210 genes expressed over 2-fold and 170 genes expressed

under 2-fold in selected AR-positive cells (n=4) vs. negative cells (n=8) (Figure 4.1C), along

with 295 genes expressed over 2-fold and 245 genes expressed under 2-fold for ERα-positive

(n=5) and negative cells (n=23) (Figure 4.2C). Consistent with QPCR data, this analysis

includes AR expression with over 2.6-fold in AR-positive cells ($p < 0.05$) and ERα with 3.3

fold higher expression in ERα-positive cells ($p < 0.05$).

Along with the survey of genetic signatures relative to steroid receptor expression, we

executed similar array-based approaches for PPARγ expression that showed an exclusive

pattern of expression in NSCLC compared to the other cancer types. We identified 633 genes

showing over 2-fold expression and 377 genes with 2-fold underexpression in lung cancer

cell lines (n=5) highly expressing PPARγ vs. cell lines with low expression (n=5) (Figure 4.3

inset). Consistently, PPARγ expression is increased more than 27-fold in the high expression

group compared to low expression group.

Moreover, genetic signatures were sorted according to Pearson correlation for the

receptors, AR, ERα and PPARγ, to closely group genes showing a similar expression pattern

to the receptor across the panel of cell lines tested. In a further test for statistical significance

($p < 0.05$) of the Pearson correlation, 195 genes (51 positively and 144 negatively correlated)

out of 380 genes and 145 genes (75 positively and 70 negatively correlated) out of 539 genes

were filtered for AR and ERα, respectively. Similarly, a total of 895 genes (568 positively

and 327 negatively correlated) out of 1010 genes were filtered for PPARγ. It is reasonable to

think that the genes statistically filtered are potentially under common control, even possibly

direct target genes, of the nuclear receptor, suggesting the idea that this type of analysis

potentially reveals a previously unidentified target gene or group of genes important in lung

cancer pathogenesis. Despite the lack of known target genes for each receptor in lung tissues,

this data interpretation is able to give insight into understanding lung cancer as a network of

genes under the control of or in coordination by nuclear receptors.


4.2.5   Genetic Signature Dependent on PPARγ Receptor and Ligand

Since the same response of cells and xenografted tumor growth in response to ligand

treatment have been attributed to receptor-dependence, the next step of my research was to

identify genetic signatures potentially responsible for the receptor-ligand-mediated tumor growth inhibition. Thus, I performed Affymetrix microarray experiments in two groups of PPARγ negative H1770 cells and positive H2347 cells. The treatment groups included vehicle control and 20hr and 48hr treatment with 3μM troglitazone (Figure 4.5). Three subsets of genes, showing both receptor and ligand-dependency, were identified by exclusion analysis that removes genes changing in a non-receptor and/or a non-ligand dependent manner. In addition, the list of genes was further sorted in a time dependent manner, 24hr and 48hr, to identify genes according to response patterns which are graphed on Figure 4.5. This analysis revealed a total of 2568 genes that are further subdivided into 1117 showing dramatically increased or decreased levels 20hr after treatment, 1026 genes continuously increased or decreased by 48hrs, and 425 genes commonly increased or decreased at both time points. I believe that further pathway analysis on a basis of molecular and biological annotations could reveal a sequence of genetic changes dependent on both PPARγ and its cognate ligand in lung cancer cells.

4.3   Discussion

Along with the validation of the diagnostic potential of the NR atlas, I further assayed cellular growth responses for a subset of NRs (AR, ERα, and PPARγ) to demonstrate the feasible utilization of NRs as therapeutic targets. Not surprisingly, only the cell lines expressing the receptor responded to cognate ligand treatment in cell proliferation assays. In addition, the observation of estrogen receptor-dependent growth inhibition upon treatment with ICI 182,780 provides pharmacological evidence that the quantitative mRNA profile of

NRs reflects functional activity of receptors selected in this study. Surprisingly, in the functional evaluation of lung cancer cell lines relative to steroid receptors with agonist treatment, I observed cell proliferation as well as growth inhibition (e.g. H1993 for AR and HCC78 for ERα) in a receptor-dependent manner, which was an unanticipated result. Consistently, several independent studies showed androgen-induced growth inhibition in AR positive cells derived from LnCaP, a well characterized prostate cancer cell for growth stimulation with DHT treatment, as well as in mouse xenograft model with the variant clones (Joly-Pharaboz et al., 2000; Umekita et al., 1996). Likewise, Osipo *et al.* showed intriguing observations in which several MCF-7 clones resistant to tamoxifen or raloxifene were sensitive to estradiol treatment in *in vitro* cell growth assays and mouse xenograft models (Liu et al., 2003; Osipo et al., 2003).

Together with pharmacological validation of the therapeutic potential of steroid receptors, lung cancer cell lines highly expressing PPARγ showed inhibition of cell proliferation in a receptor-dependent manner in the presence of the ligand troglitazone. Consistently, several lines of evidence have suggested the therapeutic potential of PPARγ in various types of cancers including breast, colon, and lung cancer (Allred and Kilgore, 2005; Bren-Mattison et al., 2005; Keshamouni et al., 2004; Sasaki et al., 2002). Furthermore, I also observed decreased cyclin D1 expression in a PPARγ-dependent manner (Figure 4.4C), which is also observed in human renal carcinoma cell lines (Yang et al., 2005). *In vivo* treatment of tumors xenografted in nude mice further demonstrated that endogenous PPARγ expression could be a therapeutic target for lung cancer treatment. Interestingly, the observation of tumor promoting effect of both PPARγ ligands, i.e. *in vitro* cell proliferation

of H1770 with troglitazone treatment and induced growth of H1770 tumor with pioglitazone

treatment in the xenograft, suggests significance of receptor-targeted therapy (Figure 4.4).

Furthermore, the previous observation of PPARγ having tumor promoting effects in breast

and prostate cancer mouse models suggests the therapeutic approaches using ligands may

need to be assessed in a tissue-specific manner (Saez et al., 2003; Saez et al., 2004). Overall,

these preclinical therapeutic evaluations using ligands for AR, ERα, and PPARγ, which have

been widely applied in the clinic, strongly support the feasibility of using the approach for

whole subclasses of nuclear receptors, rather than a single nuclear receptor, as molecular

targets for therapeutic intervention to treat lung cancer.

In analyzing the microarray data, groups of genes were listed showing more than 2-

fold higher expression in receptor-positive cells vs. negative cells. Further analysis for

statistical significance eventually revealed 195 genes for AR, 145 genes for ERα, and 895

genes for PPARγ. These genes are potentially under the same coordinated control by their

respective receptor. It is interesting to note that three orphan nuclear receptors, NGFIB,

NURR1, and COUP-TFβ were included with an average of 4.5-, 4.2-, and 3.3-fold

overexpression, respectively, in AR-positive cells. It is reasonable to interpret a pro-

inflammatory role of NR4A group. Thus, NGFIB and NURR1 may contribute to the

maintenance of lung cancer cell features along with increased AR expression. Upon

considering COUP-TFβ as a generic transcriptional repressor, it will be interesting to identify

target genes that could be potential tumor suppressors. Similarly, NOX5 and RAD18 were

identified showing negative correlations, -0.83 and -0.87 with $P < 0.05$ respectively, with 9-

fold and 2.2-fold decreased expression in AR-positive cells. NOX5 is a subtype of NADPH

oxidase, which generates reactive oxygen species (ROS) which are believed to contribute to tumor cell proliferation. Some reports describe that hypoxic conditions in tumor microenvironment activates NADPH oxidases and androgen receptor activity, suggesting that coordinated regulation of both AR and NADPH oxidase activity may be critical for tumorigenesis (De Servi et al., 2005; Goyal et al., 2004; Mehta and Mehta, 2002; Nabha et al., 2005; Park et al., 2006). Further statistical significance test ($P < 0.05$) of the Pearson correlation revealed 75 genes positively- and 70 genes negatively-correlated to similarity of ER$\alpha$ expression pattern. Interestingly, PKC delta shows a 0.62 Pearson correlation to ER$\alpha$ with 2.7-fold overexpression in the receptor positive cells. PKC$\delta$ has been shown to be potentially involved in estrogen receptor $\alpha$ mediated cell proliferation in breast cancer, implicating coordinated regulation of PKC$\delta$ and ER$\alpha$ for tumor cell proliferation even though there is no clear evidence if PKC$\delta$ is a direct target of ER$\alpha$ (De Servi et al., 2005; Nabha et al., 2005).

In addition, increased expression of VDR by 2.6-fold in ER$\alpha$-positive cells has been shown to be significant in expression patterns across the cell lines tested. Although a functional relationship of ER$\alpha$ to VDR still remains to be understood, it was proposed that co-expression of VDR with ER$\alpha$ shows better prognosis with VDR ligand treatment in breast cancer (Mehta and Mehta, 2002).

To identify groups of genes that manifest related cellular responses to troglitazone treatment, I also sorted genetic signatures of 1010 genes showing over 2-fold difference in their expression according to Pearson correlation to PPAR$\gamma$. In this analysis, several subsets of genes potentially relevant to distinct signaling pathways were identified as follows:

transforming growth factor (TGF), Wnt, and insulin-like growth factor binding proteins. For example, TGF beta 1 (35-fold), TGFβRII (3-fold), TGFβ2 (8-fold), SMAD3 (3-fold), SMAD5 (2.1-fold) for TGFβ signaling and Dickkopf homolog 1 (DKK1, 28-fold), glycogen synthase kinase 3 beta (GSK3β, 2-fold), adenomatosis polyposis coli 2 (APC2, 2-fold) for the wnt signaling pathway showed increased expression with a positive correlation to PPARγ expression. The connection of PPARγ to both pathways was previously observed in an adipogenesis model and a colon cancer model (Choy and Derynck, 2003; Jansson et al., 2005; Lee et al., 2005a; Lee et al., 2004b; Lee et al., 2005b). This observation implies that alteration of coordinated regulation of these pathways in epithelial cells could be involved in disease pathogenesis. Also, insulin-like growth factor binding proteins including IGFBP3 (43-fold), IGFBP6 (4-fold), and IGFBP7 (101-fold) were also highly expressed along with PPARγ. IGFBPs, especially IGFBP3, are associated with a good prognosis for several types of cancer including lung cancer (Lee et al., 2004b). Interestingly, *Hong et al*. reported that the ratio of serum IGF-1/IGFBP3 was significantly lower in 84 patient samples after 3 months treatment with 9-*cis*-retinoic acid, potentially due to higher level of IGFBP3 expression accompanying a lower level of IGF-1 (Lee et al, 2005a; Lee et al., 2005b). These clinical data strongly suggest that combined treatment with PPARγ ligand and 9-*cis*-retinoic acid would be a valuable chemopreventive application to the clinic. It will be interesting to analyze the physiologic relationship between PPARγ, TGFβ signaling and IGFBP signaling.

In addition, the analysis of microarray data from two lung cell lines, H1770 (PPARγ negative) and H2247 (PPARγ positive), with or without troglitazone treatment for 20hr or 48hr revealed a large number of genes such as RARγ, BUB1, SMAD3, Caspase8, NcoR1,

NR2F6, SCD, and PDK4 that showed increased expression with PPARγ ligand treatment. A direct repeat 1 (DR1) cis-element is found in the promoter of PDK4 which is believed to be potential target of PPARδ in liver and heart. Interestingly, PDK4 was induced 4-fold at both 20hrs and 48hrs after the PPARγ agonist troglitazone treatment in H2347 cells but not in H1770, which is PPARγ negative. Further, RARγ and NR2F6 were also induced about 20-fold and 5-fold, respectively, in H2347 cells treated with the ligands at both time points. It will be of interest to investigate the biological relevance of those receptors in PPARγ-dependent tumor cell growth inhibition.

Other genes also are intriguing; for instance, SCD and PDK4 are known to be involved in both fatty acid and glucose metabolism. The genes down-regulated by the ligand treatment may be due to the induction of NcoR1 and NR2F6 both of which are known to repress gene expression. Thus, it will be crucial to determine if PPARγ bound with ligand is actively involved in repression of certain tumorigenic genes. A recent publication shows that sumoylation of PPARγ by rosiglitazone treatment actively represses iNOS mRNA expression by stabilizing corepressor complexes on the iNOS promoter. This provides insights to how PPAR may repress genes in lung cancer (Pascual et al., 2005). Thus, further molecular approaches, gain-of- and loss-of-function studies, for genes believed to be potentially regulated by PPARγ will provide a more detailed molecular pathway for disease progression. Moreover, surveys of global genetic signatures with microarray experiments in H1770 and H2347 tumors with or without pioglitazone treatment can be compared to the genetic signatures obtained from Figure 4.5, which will provide an insight into the systemic effects of drug treatment as well as the tumor microenvironment *in vivo*.
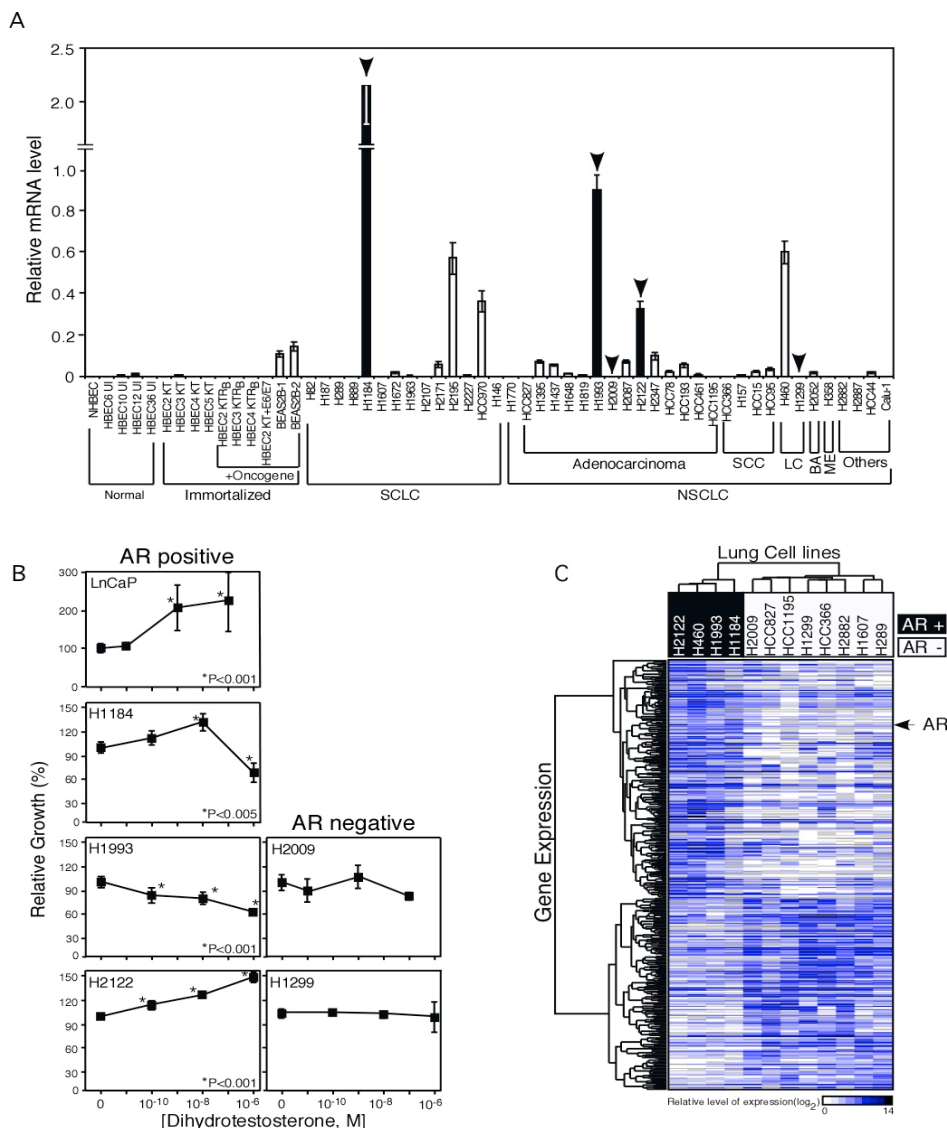
Figure 4.1 Pharmacological evaluation of androgen receptor in lung cancer cells.
(A) Expression profile of androgen receptor in a panel of lung cells. Arrowheads indicate cell lines (black bar) selected for AR agonist DHT treatment.
(B) Evaluation of AR for cell growth responses to DHT treatment. Relative % growth response was assessed as described in methods. *ANOVA* was performed for statistical analysis for the growth responses. The asterisks show statistical significance.
(C) Unsupervised clustering analysis of both genetic signatures and lung cancer cells. A genetic signature of 380 genes showing more than 2-fold difference between AR positive and negative lung cancer cells was applied for unsupervised clustering analysis using Matrix 1.29 as described in methods. Groups represent AR-positive cells with black box and AR-negative cells with white box. An arrow indicates position of androgen receptor.
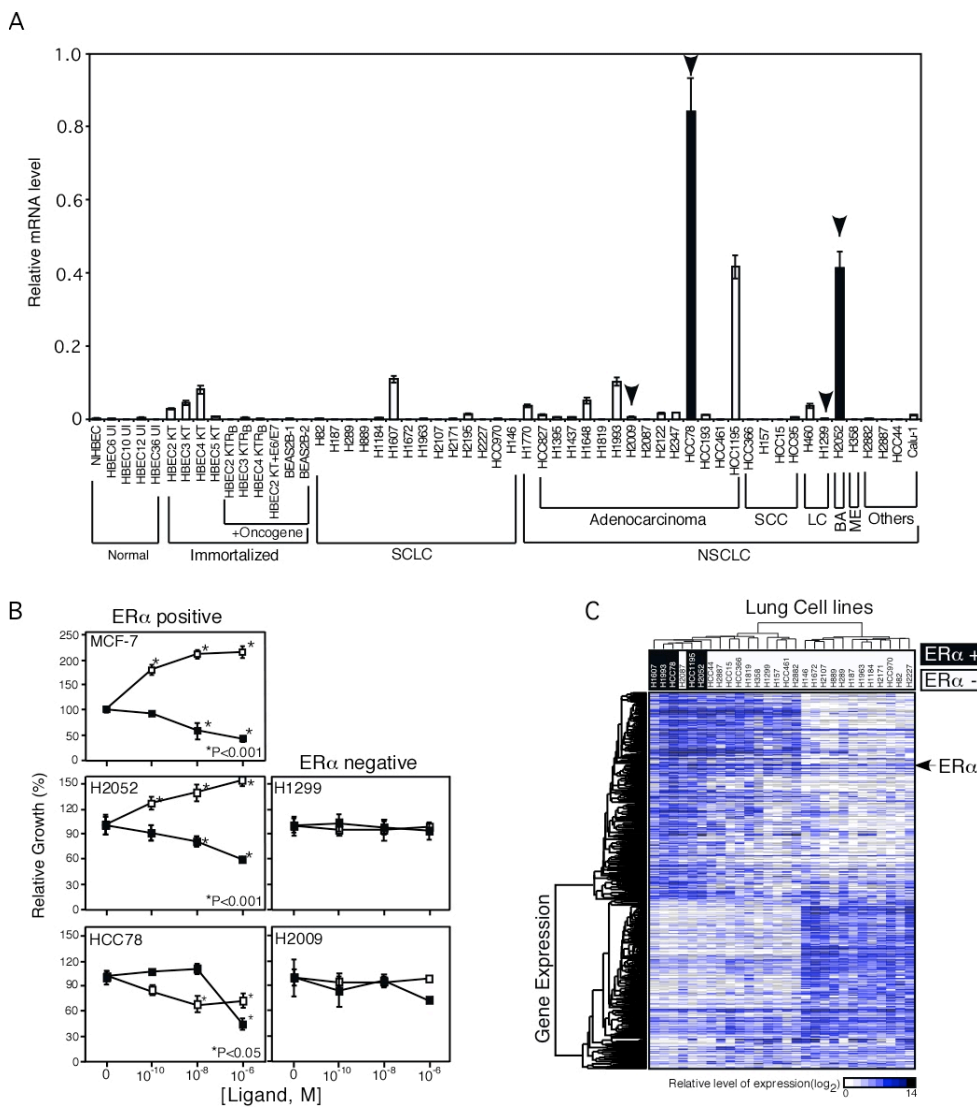
Figure 4.2 Pharmacological evaluation of estrogen receptor α in lung cancer cells.
(A) Expression profile of estrogen receptor α in a panel of lung cells. Arrowheads indicate cell lines (black bar) selected for treatment of ERα ligands.
(B) Evaluation of ERα for cell growth responses to ligands treatment. Relative % growth response was obtained as described in methods. The X-axis represents concentration of ligands treated. The open square (□) and closed square (■) represent growth response to agonist 17β-estradiol and antagonist ICI 182, 780, respectively. *ANOVA* was performed for statistical analysis for the growth responses. The asterisks show statistical significance.
(C) Unsupervised clustering analysis for both genetic signatures and lung cancer cells. A genetic signature of total 540 genes showing more than 2-fold difference between ERα-positive and negative lung cancer cells was performed using same application as AR in figure 4. Groups highlighted represent ERα-positive cells with black box and ERα-negative cells with white. An arrow indicates position of estrogen receptor α.
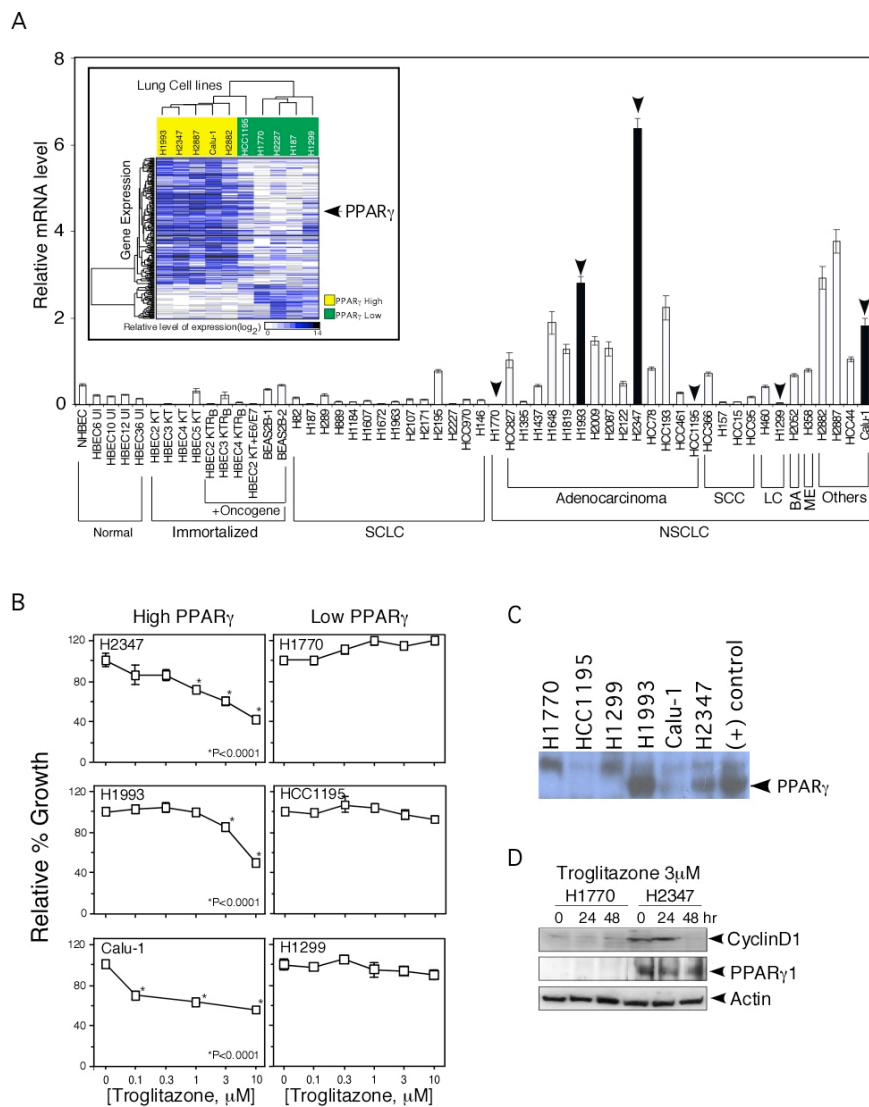
Figure 4.3 Pharmacological evaluation of PPARγ in lung cancer cells.
(A) Expression profile of PPARγ in a panel of lung cells. Arrowheads indicate cell lines (black bar) selected for treatment with PPARγ ligand troglitazone. **(Inset)** Unsupervised clustering analysis was performed for both genetic signatures and lung cancer cells. A genetic signature of total 1010 genes showing more than 2-fold difference between PPARγ-positive and negative lung cancer cells was performed using same application as steroid receptors in figure 4 and 5. Groups highlighted represent PPARγ-positive cells with black box and PPARγ-negative cells with white box. Each arrowhead indicates the selected lung cell lines according to the PPARγ expression.
(B) Evaluation of PPARγ for cell growth responses to troglitazone treatment. Relative % growth response was obtained using MTT assay as described in methods. The X-axis represents concentration of troglitazone. The asterisks show statistical significance.
(C) Western blot analysis was performed for PPARγ expression lung cancer cells selected. Cell lysates from differentiated 3T3L-1 adipocyte were used as a positive control for the receptor detection. (D) Cyclin D1 expression with troglitazone treatment. Western blot analysis was performed for cyclin D1, PPARγ1 and actin for sample loading as described in methods.
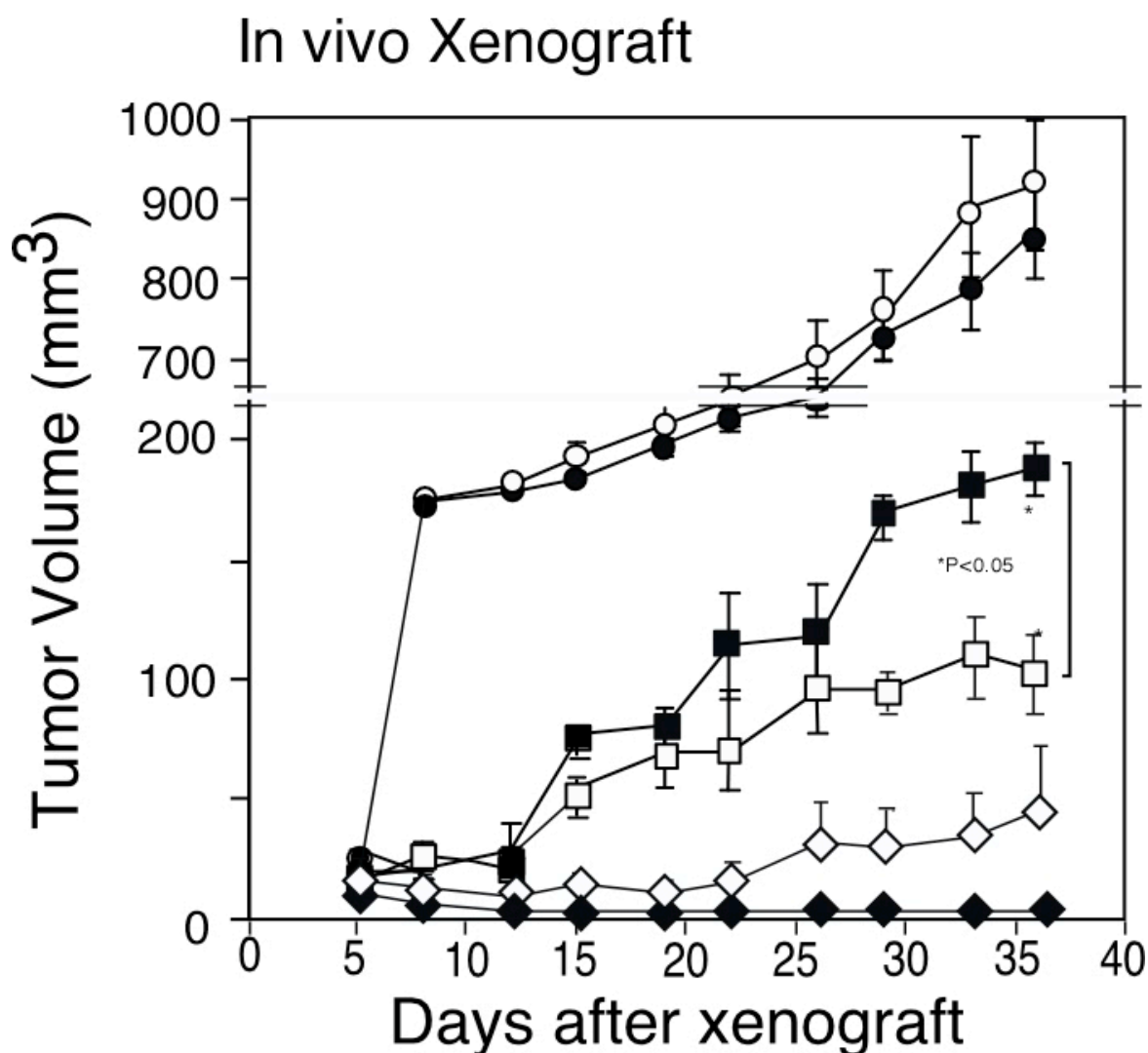
Figure 4.4 *In vivo* treatment of xenograft tumors with the PPARγ ligand pioglitazone. The H1770, H1299 and H2347 ($2X10^6$) cells were injected into the right flank of athymic nude mice, followed by the treatment with 25 mg/kg pioglitazone or vehicle control four times a week; open circle (O), H1299 treated with the pioglitazone; filled circle (●), H1299 treated with the vehicle; open square (□), H2347 treated with the pioglitazone; filled square (■), H2347 treated with the vehicle; open diamond(◊), H1770 treated with the pioglitazone; filled diamond (◆), H1770 treated with the vehicle. The tumor volumes were measured twice a week. The values represent mean tumor size (*n* = 4 per group) with standard error. Statistical analysis with student *t*-test using SPSS 11.5 software (SPSS Inc., Chicago, IL) revealed significant delay of tumor growth with pioglitazone treatment of H2347 tumor expressing PPARγ (*P* < 0.05).

A

A  Affymetrix chip expts (6 conditions)

H1770 (PPARγ negative)            H2347(PPARγ positive)

Troglitazone        Veh    20hr    48hr            Veh    20hr    48hr

Genes showing
change (2 fold) in H1770
with Troglitazone
treatment

Total genes ~ 40,000 genes

Exclude genes showing under 2 fold difference
in both 20hr and 48hr treatment
compared to vehicle treatment

20hr                                    48hr

Up 914
(12 down in 20hr)

1212 Down 598    2210          2799    Up 889

Up 614          Down 1296              Down 1910
                (104 up in 20hr)

Up
RARγ, BUB1, SMAD3, Caspase8, NcoR1, FOXOa3, NR2F6, SCD, PDK4
Down
MALAT-1 (metastasis associated Lung adenocarcinoma transcript 1(non-coding RNA)

B



Figure 4.5 Identification of potential target genes of PPARγ in lung cancer cell lines. (A) A schematic representation on microarray experiment and procedure (in *venn diagram*) identifying potential target genes from the bioinformatics analysis of data. (Bottom) A list is shown for genes that are up-regulated or down-regulated with PPARγ activation.
(B) Expression pattern of individual genes regulated by PPARγ activation in a time course after troglitazone treatment. The individual genes are plotted in a time-course (0hr, 20hr, 48hr). The x-axis represents time after treatment with troglitazone and the y-axis stands for expression level normalized by the vehicle treatment in $\log_2$ ratio.

# CHAPTER FIVE
## Understanding Lung Cancer Pathogenesis from the
## Viewpoint of Nuclear Receptors

### 5.1 Introduction

The nuclear receptor (NR) superfamily comprises 48 members, mostly ligand-activated transcription factors, which share similar structures including two activation domains, AF1 and AF2, and a DNA binding domain (DBD) (Chawla et al., 2001). Although endogenous or natural ligands are not known for a subset of these transcription factors, called orphan receptors, extensive studies over the last two decades show their significance in development (Gu et al., 2005; Lu et al., 1997; Morriss-Kay and Ward, 1999), differentiation (Fu et al., 2005; Lee et al., 2004a) and physiology (Akashi and Takumi, 2005; Chawla et al., 2001; Yin and Lazar, 2005; Yin et al., 2006). A dysregulation of NR pathways causes severe chronic diseases such as diabetes (Blaschke et al., 2006a; Ramachandran et al., 2006), atherosclerosis (Barish, 2006; Blaschke et al., 2006b; Lee and Plutzky, 2006; Moreno et al., 2006; Rizzo and Fiorucci, 2006), and several types of cancer (Abu et al., 2005; Allred and Kilgore, 2005; Benner et al., 1995; Bren-Mattison et al., 2005; Mehta and Mehta, 2002; Nakagawa et al., 2005a; Wang et al., 2006d). Several lines of genetic and epigenetic evidence show that lung cancer incidence is increased by direct or second-hand smoke (Gazdar and Carbone, 2003; Minna et al., 2003; Zhu et al., 2003). Genetic screening in tumor specimens from lung cancer patients shows that genetic mutations occur in 30% of non-small

cell lung carcinomas (NSCLC) in the *K-ras* gene and 50% of small cell lung carcinomas (SCLC) in the p53 gene. Transgenic and knock-out mouse models for both oncogenes have demonstrated that the genetic mutations observed in lung cancer patients are critical for cancer progression (Johnson et al., 2001; Meuwissen et al., 2003). In addition, promoter silencing by methylation of the p16$^{INK4a}$ cell cycle regulator gene is found in 67% of adenocarcinomas (ADK) and 70% of squamous cell carcinomas (SCC) of lung tumors (Jarmalaite et al., 2003; Kim et al., 2001; Minna, 2005). Furthermore, recent studies identified additional chromosomal hot spots such as 5q, 11q23.2, along with the well-known tumor suppressor locus 3p21.3, which show genetic and epigenetic alterations in high-risk groups or cancer patients (Belinsky, 1998; Minna, 1993). Although this prevalence of genetic and epigenetic alterations is critical in lung cancer pathogenesis, the molecular mechanisms involved remain to be elucidated.

A number of studies have shown the relevance of individual NRs to the onset, development, treatment, and chemoprevention of cancer. For instance, overexpression of the retinoic acid receptor alpha (RARα) due to a fusion with PML (RARα/PML) and estrogen receptor alpha (ERα) expression causes the onset of leukemia and the progression of breast cancer, respectively (Jordan, 2002; Jordan, 2004; Wang et al., 2006b; Yoo et al., 2006). Targeting ERα using the selective modulators tamoxifen or raloxifene and blockade or ablation of dihydrotestosterone (DHT), which is the strongest endogenous ligand for androgen receptor (AR), are well-known therapeutic strategies to treat breast and prostate cancers, respectively. In addition, retinoic acid, the ligand for retinoic acid receptors, has been proposed as a potential drug for chemoprevention for lung cancer which has the highest

incidence amongst all forms of cancers (Khuri and Lippman, 2000; Lippman et al., 1993).

Since multiple NRs have potential relevance to cancer biology in a variety of tumors, I

wanted to investigate how the NR superfamily is involved in lung tumor pathogenesis using

*in vitro* and *in vivo* lung cancer models as depicted in section 2.2. NRs have distinct

expression patterns, either overexpressed or underexpressed in the HBEC3 panel, caused by

oncogenic manipulations. Further biological approaches will be discussed regarding the

potential relevance of NRs in control of inflammatory pathways, which are widely believed

to cause various chronic diseases, especially cancer incidence, in the following sections.

**5.2 Results**

5.2.1 Characterization of Immortalized Cells

To identify the tumorigenic potential of HBEC3 by oncogenic alterations, a series of

daughter cell lines were established with the following alterations 1) *K-ras$^{V12}$*, 2) knockdown

of p53, or 3) both changes together. The stable knockdown of p53 was confirmed for both

mRNA and protein expression using quantitative real-time PCR and immunoblot analysis,

respectively. In addition, the activity of stably introduced oncogenic *K-ras$^{V12}$* was assessed

by measuring phosphorylation of MEK, a downstream target kinase of *K-ras* (Figure 5.1).

More interestingly, these genetic changes clearly induced vacuole-like cellular morphological

changes (Fig 5.1B). These data suggest that physiological changes have occurred in these

cells.

5.2.2 Identification of NRs with Potential Relevance to Lung Cancer Pathogenesis

To explore the physiological relevance of NRs in lung cancer pathogenesis, the mRNA expression level of all 48 members of the superfamily was profiled using quantitative real-time PCR in the genetically defined panel of immortalized human bronchial epithelial cells (HBEC3), as mentioned in section 2.2. Comparative analysis of the expression profile revealed that several NRs, including COUP-TFI, ERα, NGFIB3, PPARα, PPARδ, PPARδ2, PPARγ RARβ, Rev-erbα, RORα, and RXRβ, showed a distinct expression patterns across these cell lines (Figure 5.2). The introduction of oncogenic $K$-$ras^{V12}$ in either a WT or a p53 knock-down background resulted in increased or decreased expression of the above receptors (Figure 5.2). Interestingly, all PPARs, except PPARγ2, were increased, strongly implicating that further post-immortalization processes, induced by oncogenic $K$-$ras^{V12}$, may accompany changes in fatty acid metabolism. This profile strongly supports the notion that NR expression is involved in disease progression by both of these well-known oncogenes.

Further analysis of the NR atlas in tumorigenic subclones revealed different patterns of NRs except COUP-TFI, RARβ, and PPARγ that are consistently induced in both non-tumorigenic parental culture and tumorigenic clones. These data suggest that a sequential change of multiple NRs is necessary on transformation from an immortalized non-tumorigenic stage to a tumorigenic stage. For instance, RARβ expression was expressed at relatively lower level in the non-tumorigenic parental stages but rebounded up to more than 20-fold in the C5 tumor grown as a xenograft when normalized to HBEC3 immortalized with CDK4 and hTERt (Figure 5.2). Based on the evidence in the literature that RARβ expression is associated with anti-tumorigenic properties, it will be interesting to investigate whether or not the tumor shows differential response to the RAR agonist in tumorigenic progression

using this panel of immortalized HBECs (Toulouse et al., 2000; Zhu et al., 1996). On the other hand, the nuclear receptor COUP-TF I, generally known as repressor, displayed a continuous increase of expression when cells progressed from immortalized through tumorigenic stages. Thus, the next approach is to define the role of COUP-TFI by performing a variety of functional studies such as loss-of-function (e.g., siRNA) and gain-of-function approaches at both the *in vitro* cellular level and the *in vivo* xenograft level of tumorigenesis. PPARγ is of particular interest because this receptor has been extensively studied as a therapeutic target in various types of cancers. In addition, VDR protein showed decreased expression despite minor changes at the mRNA level, suggesting that posttranslational stability of this particular receptor is regulated by the combination of oncogenic introductions (Figure 5.5).

5.2.3 PPARγ and VDR in Lung Cancer

In this chapter, I mainly focus on anti-inflammatory molecular mechanisms of PPARγ which contributes to the anti-tumorigenic process, and show preliminary data for VDR. Consistent with the previous reports showing that COX2 expression is involved in inflammation as well as tumor progression (Lynch and Kim, 2005; Prueitt et al., 2007; Rodrigues et al., 2004; Singh et al., 2005; Witton et al., 2004), I observed dramatically increased expression of COX2 in the HBEC3 cells with oncogenic *K-ras*$^{V12}$ (Figure 5.3). Interestingly, COX2 expression is coincident with PPARγ expression in HBEC3, as well as a panel of lung cells, strongly suggesting the relevance of the two proteins in disease progression. Given that the transcription factor PPARγ plays a significant role in anti-

inflammation, it is conceivable that PPARγ activation possibly inhibits pro-inflammatory COX2 expression. Interestingly, troglitazone treatment inhibited expression of both COX2 mRNA and protein. Expression of iNOS, another well-known pro-inflammatory factor, is negatively regulated by sumoylated PPARγ which stabilizes the corepressor complex (Pascual et al., 2005). Furthermore, cyclin D1 expression is decreased with troglitazone treatment only in HBEC3 cells with the dual oncogenic alterations, which is consistent with previous observations (Shao et al., 2002; Yang et al., 2005). These data support the idea that PPARγ potentially inhibits tumorigenesis by inhibiting COX-2 and cyclin D1 expression (Figure 5.3 bottom).

To see if decreased expression of cyclin D1 by troglitazone treatment affects cell cycle progression, I assayed HBEC3 cell proliferation using the MTT assay. Interestingly, troglitazone treatment does not inhibit cell proliferation, even with the selective COX2 inhibitor celecoxib (Figure 5.4). In addition, the observation of decreased expression of VDR consistently reflects a decreased response to 1,25α dihydroxyvitamine–D3 treatment in HBEC3 KTR$_L$p53-/- cells, suggesting a receptor-dependent growth inhibition with cognate ligand treatment (Figure 5.5). Interestingly, treatment with VDR ligand in the presence of 1μM of PPARγ agonist troglitazone sensitized both HBEC3s expressing oncogenic $K\text{-}ras^{V12}$ (HBEC3-KTR and HBEC3-KTR$_L$p53$^{-/-}$), supporting the above idea that PPARγ activation somehow counteracts the $K\text{-}ras$ pathway (Figure 5.5). This sensitization seems not to be directly relevant to VDR because VDR expression is still low in the presence of troglitazone (Figure 5.5). Thus, it may be interesting to determine the physiological interplay between VDR and PPARγ in tumors.

5.2.4 PPARγ expression is lost in tumorigenic clones of HBEC3 cells

Based on the result that both PPARγ and COX2 expressions were dramatically increased in both HBEC3s expressing oncogenic *K-ras$^{V12}$* (HBEC3-KTR and HBEC3-KTR$_L$p53$^{-/-}$), I wanted to confirm the expression in xenograft tumors as well as in tumorigenic clones. Surprisingly, both mRNAs were reduced to similar or less levels than in HBEC3-KT, implicating that molecular dissection of the functional roles for both PPARγ and COX2 are critical to understanding lung cancer pathogenesis induced by oncogenic *K-ras*. In addition, we confirmed that both oncogenic alterations were well maintained in the immortalized cells (Figure 5.6). Interestingly, the C-5 tumorigenic clone showed even stronger oncogenic K-ras activity than the parental clone, together with a similar level of p53 knockdown to the parental clone (Figure 5.6).

5.2.5 Nuclear Receptors in Lung Cancer Mouse Model

Since the functional relevance of PPARγ with oncogenic alterations was observed in the HBEC model, along with identification of multiple NRs with manifest changes in mRNA expression, I further pursued the NR profile in a mouse lung cancer model which autonomously develops lung adenoma and adenocarcinoma due to constitutive expression of oncogenic *K-ras$^{V12}$*. The NR profile in pair-matched lung samples revealed multiple groups showing distinct expression patterns, i.e., NRs with low expression, NRs expressed but no difference between normal and corresponding tumor tissues, differentiate expression between individual mice, and different expression between normal and tumor but common across all

mice, as shown in Figure 2.7 and 2.9. Along with the NR profile in age-matched wild type mouse lung samples, unsupervised hierarchical clustering analysis was applied to the dataset, which revealed nice segregation of the samples into wild type, normal, and tumor tissues (Figure 5.7). In addition to the generic classification above, the next question was whether any single NR or subset of NRs changed as the pathogenesis progressed. Strikingly, the normal expression of a subset of NRs is already disrupted even in histologically normal tissues from the *K-ras*$^{V12}$ mouse, and shows severe loss of a cluster of NR expression (Figure 5.8). This observation occurs in both genders, implying that NR expression changes can be utilized to detect earlier preneoplastic changes which may be difficult to assess by histological examination alone (Figure 5.8).

## 5.3 Discussion

For the last two decades, since the cloning of the glucocorticoid receptor, a growing number of studies have explored significant physiological roles for nuclear receptors, which includes energy balance, circadian rhythm, lipid homeostasis, reproduction, and immune response. Abnormal regulation of NRs can cause severe chronic diseases such as obesity, diabetes, metabolic syndrome X, immune dysregulations, and cancer (Joseph et al., 2004; Li et al., 2006; Mehta and Mehta, 2002; Nakagawa et al., 2005a; Saez et al., 2004; Tobin and Freedman, 2006; Wang et al., 2006a; Wang et al., 2006c; Zelcer and Tontonoz, 2006). Also, several lines of evidence are suggestive of the diagnostic and prognostic value of various receptors along with their therapeutic potential as targets in various types of cancer. For example, immunohistologic staining for estrogen receptor and progesterone receptor has

been a routine method not only for diagnostic and prognostic purposes, but also for

determining the use og antagonists such as tamoxifen and raloxifene in breast cancer

(Aggarwal et al., 2005). Other receptors such as AR, PPARγ, and CAR, have shown

therapeutic potential in prostate cancer, colon cancer, and hepatocellular carcinoma,

respectively (Aung et al., 2006; Botrugno et al., 2004; Chmelar et al., 2007; Knapp et al.,

2006; Li et al., 2006; Schroder et al., 2006; Vandoros et al., 2006; Yamamoto et al., 2004). In

addition, recent studies have shown the involvement of the orphan receptors LRH-1 and

PPARδ in breast and colon cancer (Aung et al., 2006; Botrugno et al., 2004). Retinoic acid, a

vitamin A derivative which is a ligand for RAR and RXR, has been proposed for

chemoprevention against lung cancer (Lippman et al., 1993; Wang et al., 2006d).

Collectively, NRs play significant roles in pathogenesis and/or progression of cancer.

To demonstrate the potential role of NRs in lung cancer pathogenesis, all 48 NRs

were profiled using quantitative real-time PCR to assess mRNA levels in a panel of

immortalized human bronchial epithelial cells which harbor oncogenic alterations (e.g., $K$-$ras^{v12}$, $p53^{-/-}$, or both) as well as an *in vivo* mouse model from which tissue samples were

obtained from wild type, and from normal and corresponding tumor of the oncogenic $K$-$ras^{V12}$ mouse model in both genders. In comparative analyses of the NR profile in both

models, HBECs and the mouse model, certain discrepancies have been observed. Most

receptors except PPARγ were included either in the non-tumorigenic or tumorigenic group,

suggesting that different receptors could be differently involved in either immortalization or

the tumorigenic process. Interestingly, the pattern of PPARγ expression, which was

dramatically increased in immortalized cells, was reversed to even lower than the control

level (e.g., HBEC3-KT) in tumorigenic clones. Thus, further studies should include a

xenograft mouse model of the tumorigenic clones stably expressing an inducible PPARγ,

followed by treatment of the tumors with PPARγ agonist to demonstrate the anti-tumorigenic

function of PPARγ, as well as a potential target for chemoprevention of cancer incidence.

Second, only increased expression of LRH-1 in tumor appears in common between the

tumorigenic clones and the mouse model. Although both systems induce tumor formation

with oncogenic $K\text{-}ras^{V12}$, cellular heterogeneity of tissue from the mouse model, including

tumor microenvironmental factors and systemic effects can be attributed to this discrepancy.

Different species may have different receptor sets generating similar physiological outcomes,

even though individual receptors have similar functions in most cases. This is supported by

the recent observation comparing the NR superfamily between mouse embryonic stem cells

and human embryonic stem cells in various stages of embryoid body formation (data not

shown). The same nuclear receptor pair showed unexpectedly low Pearson correlations but

much higher correlation with other receptors, potentially suggesting that a higher network of

a subset of NRs may be critical for controlling detailed physiological responses. Thus, since

we determined what subgroups of NRs are correlated between these two species, I believe

that it is possible to predict human physiological outcomes using mouse models.

Consistent with the previous report showing that heterozygotes of LRH-1 transgenic

mice are susceptible to tumorigenesis, the increased expression LRH-1 in both tumorigenic

clones and mouse tumors normalized by corresponding normal tissues suggests a couple of

questions (Botrugno, 2004; Mueller, 2006 ; Schoonjans, 2005). First, is LRH-1 necessary to

induce lung tumorigenesis? Second, what are the potential targets of LRH-1? In analyzing

the subset of NRs, it will be interesting to uncover the biologic relation or molecular mechanism between the nuclear receptors or NR sets and to $K\text{-}ras^{V12}$, further extending to other cell survival pathways or kinase pathways (e.g. EGFR) activated in most cancers.

In further functional analysis with PPARγ and COX2 proteins, it will be interesting to perform tumor invasion assays using PPARγ agonists in combination with COX2 inhibitors. Given that both proteins are physiologically relevant to adipogenesis, it is of interest to know how tumorigenic processes are related to lipid metabolism. Overall, further biological studies of the NR subsets are needed to understand pathogenesis.

A



K-ras

pMEK

Actin

p53

QPCR

p53/18S

8

6

4

2

0

1    2    3    4

1. HBEC3 CDK4+hTERt
2. HBEC3 CDK4+hTERt+pSRZ+pLenti-K-ras$^{V12}$
3. HBEC3 CDK4+hTERt+p53$^{-/-}$
4. HBEC3 CDK4+hTERt+p53$^{-/-}$ +pLenti-K-ras$^{V12}$

B



pSRZ : pSuper-retro(zeocine) -- stable siRNA vector
pLenti: Lentiviral vector
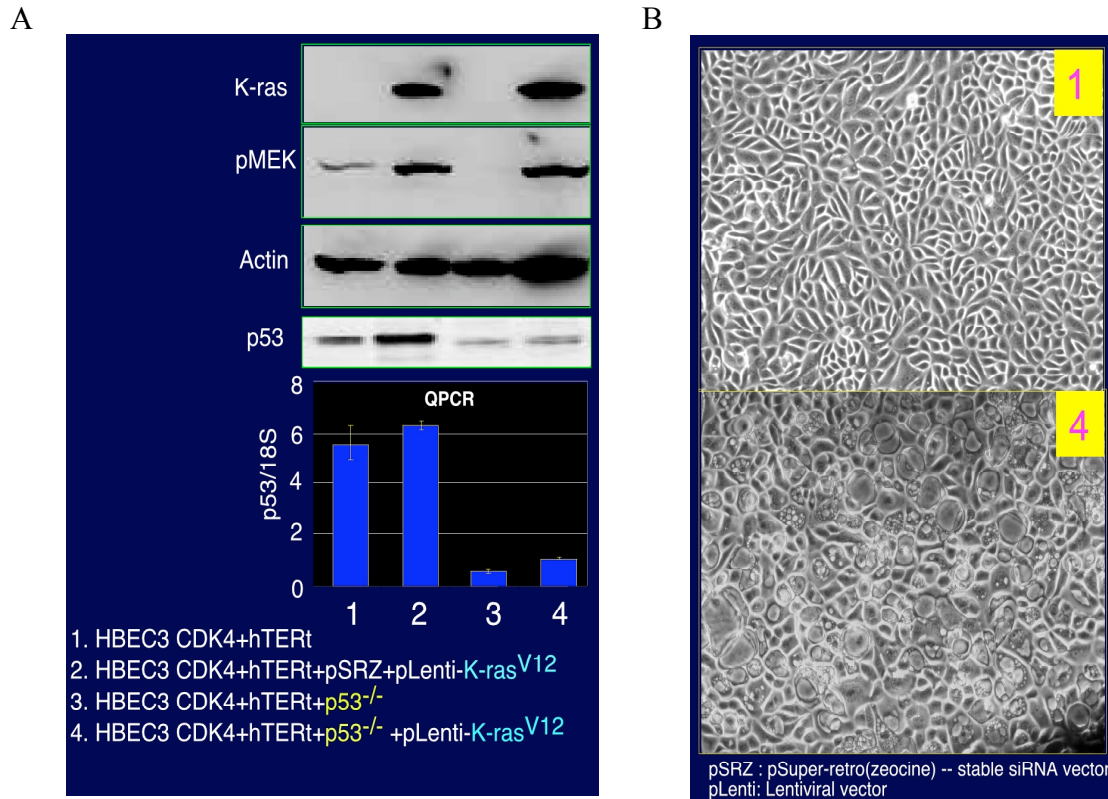
Figure 5.1 Characterization of immortalized human bronchial epithelial cells with oncogenic alterations.
(A) Western blot and quantitative real-time PCR assays. The expression of oncogenic $K$-$ras^{v12}$ is shown to be functionally active, which is shown by MEK phosphorylation. Both quantitative real-time PCR and western blot assays confirmed knockdown of p53 in immortalized cell lines with stable knockdown of p53. The real-time PCR primer for p53 is listed in table 1. Actin is used as loading control. (B) Morphological change of the immortalized cell lines harboring both knockdown of p53 and oncogenic $K$-$ras^{v12}$. Both oncogenic alterations induce cellular pathology accompanying the morphological change. As previously described, normal human bronchial epithelial cells are immortalized with CDK4 and hTERt, a catalytic subunit of human telomerase, and followed by further oncogenic alterations with either p53 knockdown by stable small interference RNA technique, introduction of oncogenic $K$-$ras^{V12}$ using lentiviral system, or both. The immortalized cells are HBEC3 CDK4+hTERt+pSRZ+pLenti-LacZ, HBEC3 CDK4+hTERt+pSRZ+K-rasv12, HBEC3 CDK4+hTERt+p53-/-+pLenti-LacZ, HBEC3 CDK4+hTERt+p53-/-+K-ras, where pSRZ and pLenti-LacZ represent pSuper-retro-zeocin vector and lentiviral vector expressing lacZ, respectively.

A
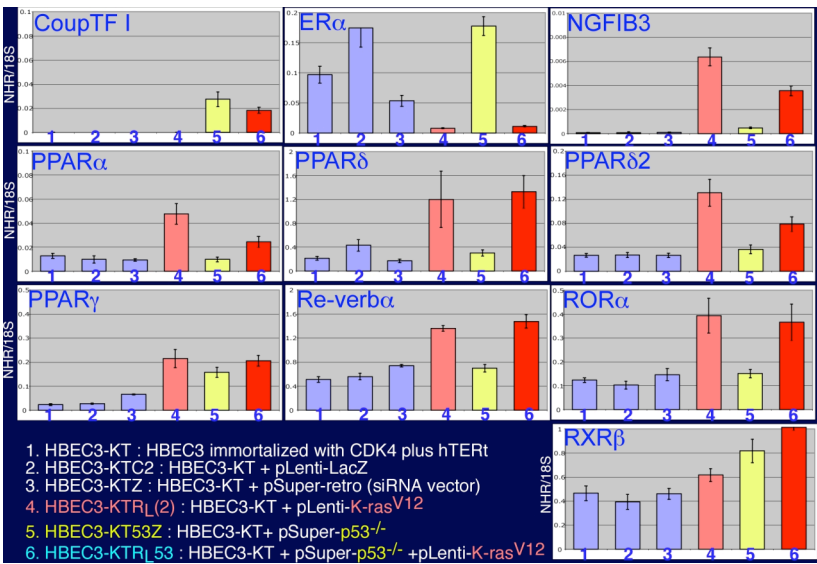


B
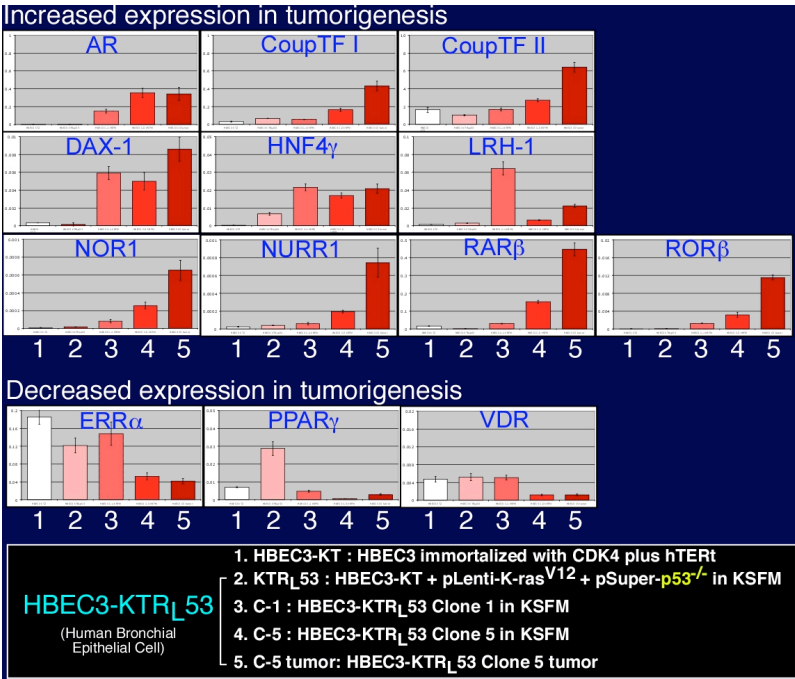


Figure 5.2 Subsets of NRs Relevant to Tumorigenesis.
(A) A subset of NRs is oncogenic dependent pattern of expression. (B) A subset of NRs is shown in a tumorigenic-specific pattern of expression.
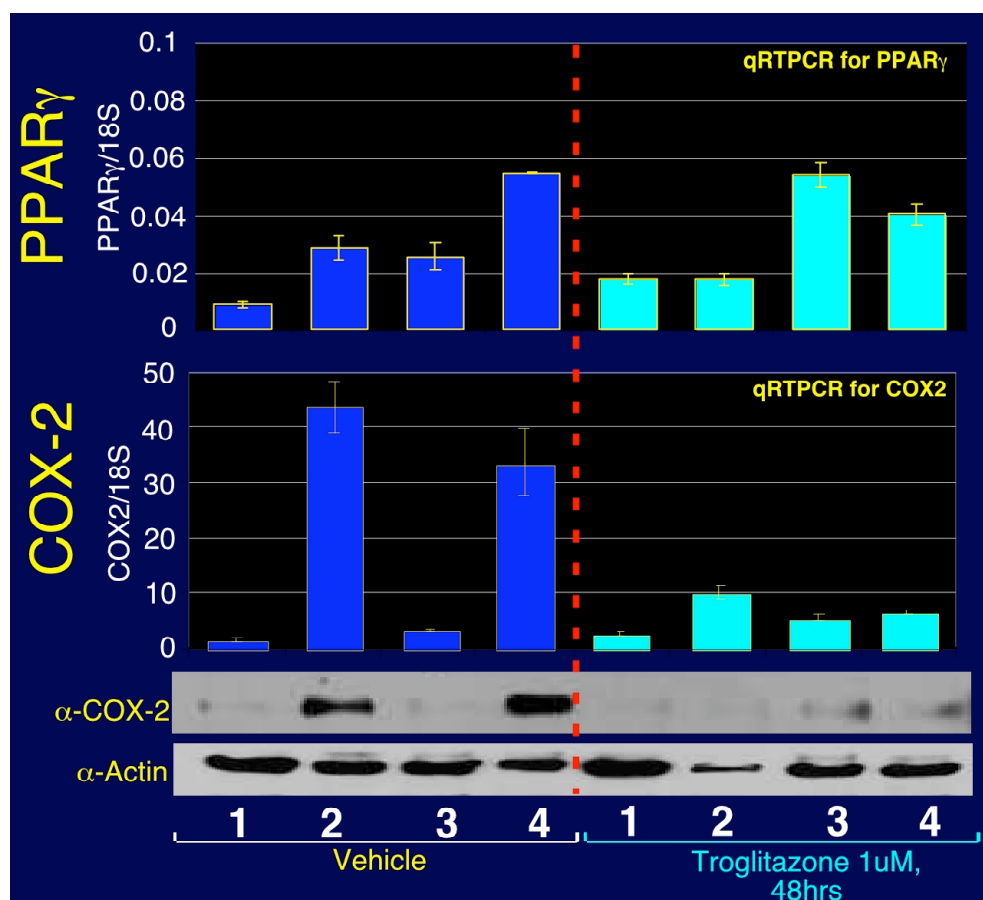
Figure 5.3 Expression of PPARγ and COX-2 in HBEC3s with various oncogenic alterations. Quantitative mRNA expressions for PPARγ and COX-2 was examined with or without treatment of 1uM troglitazone, along with COX-2 protein expression. Numbers represent HBEC3 cells with different oncogenic alterations; 1, vector control; 2, $K-ras^{V12}$; 3, p53 knocked-down; 4, $K-ras^{V12}$+ p53 knocked-down.
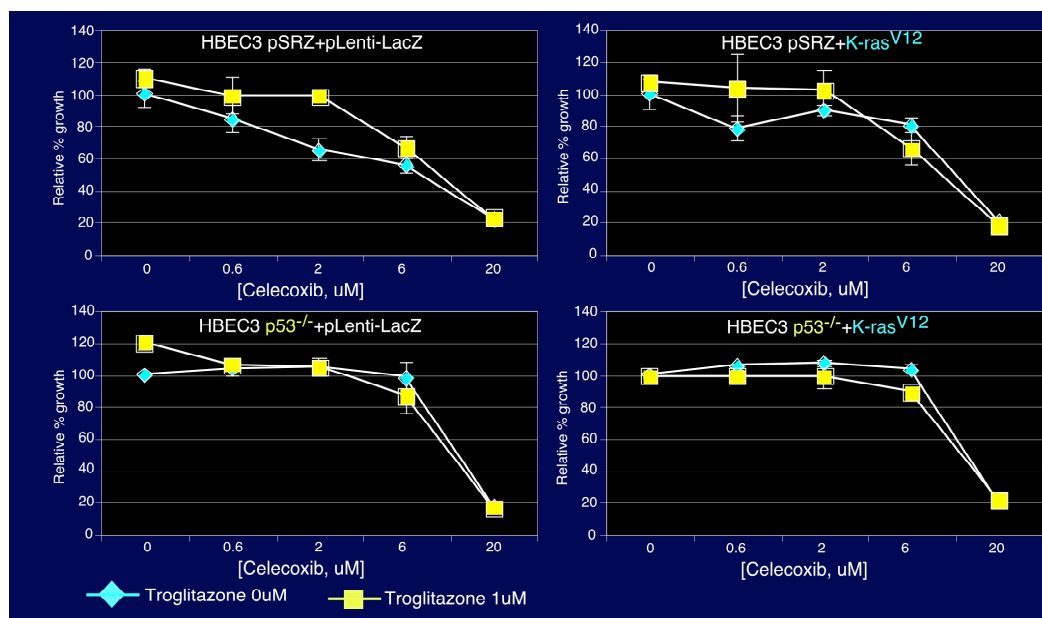
Figure 5.4 Cell Growth Response to the Treatment with Troglitazone or COX2 Inhibitor Celecoxib. Growth response of HBEC3 cells with various types of genetic modification were measured using MTT assay after treatment with celecoxib in a dose-dependent manner in the presence or absence of 1uM troglitazone.
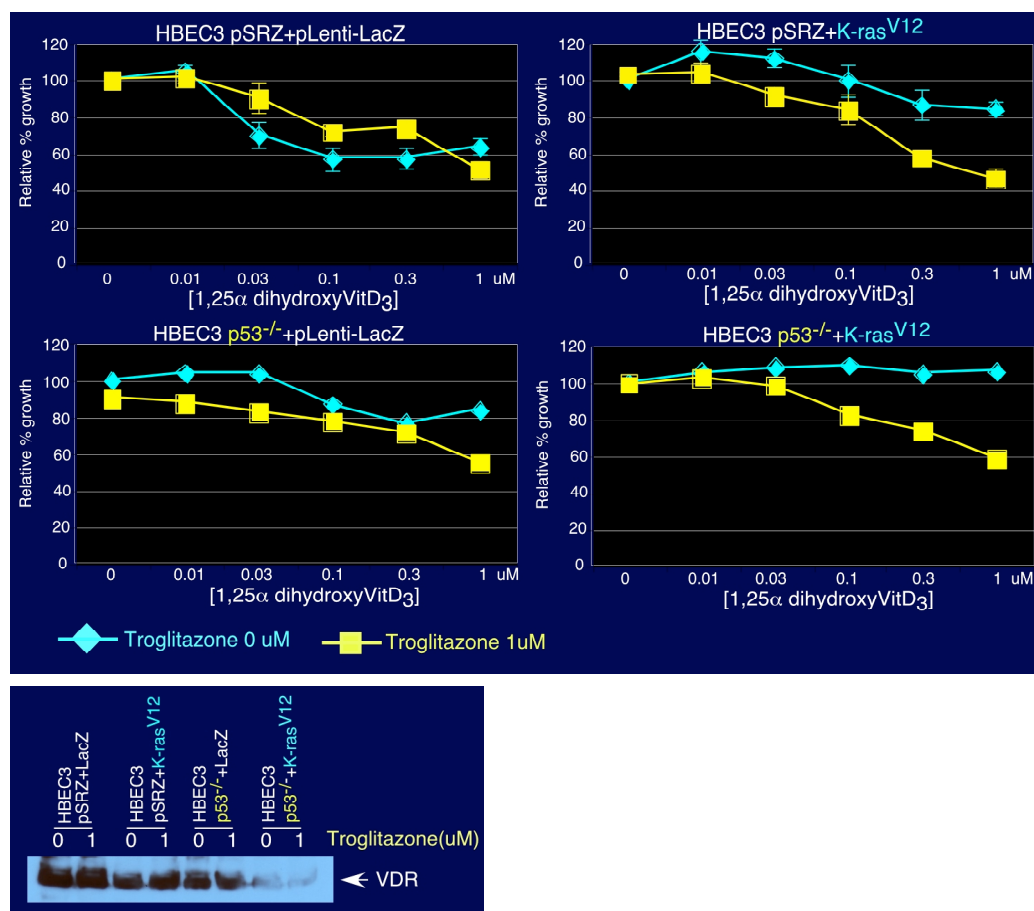
Figure 5.5 Cell Growth Responses to the Treatment of Vitamin D Receptor (VDR) Ligand (1,25α dihydroxyvitamine D3). (Top) Growth response of HBEC3 cells with various types of genetic modification were measured using MTT assay after treatment with 1,25α dihydroxyvitamine D3 in a dose-dependent manner in the presence or absence of 1uM troglitazone. (Bottom) Western blot assay shows VDR protein expression in the HBEC3s treated with vehicle or 1uM troglitazone.
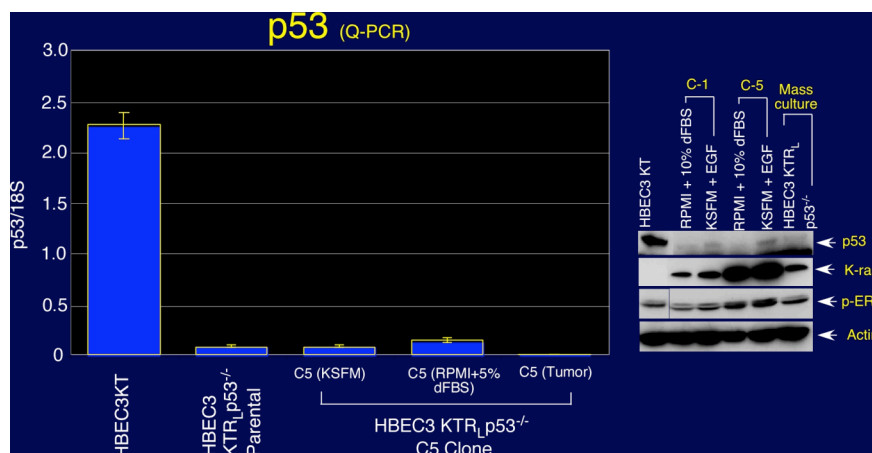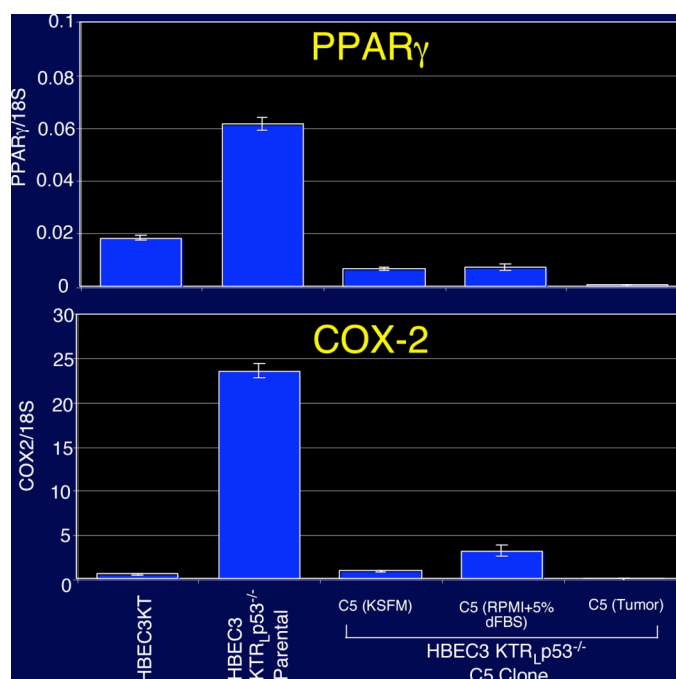
Figure 5.6 Characterization of tumorigenic clones.
(A) Identification of genetic alterations in tumorigenic C-1 and C-5 clones. Quantitative real-time PCR assay shows stable knockdown of p53 in parental cells, and cultured C-5, and tumor C-5 (left panel). Western blot analysis shows p53 knockdown, and oncogenic $K$-$ras^{V12}$ expression together with its activity by showing ERK phosphorylation in the same panel used for qRT-PCR of p53 (right panel).
(B) Identification of both PPARγ and COX-2 expression in the tumorigenic C-5 clone and xenografted tumor. Quantitative real-time PCR assay shows mRNA expression for PPARγ (upper panel) and COX-2 (lower panel)among immortalized cells without oncogenic alterations, parental mass culture with dual oncogenic alterations, tumorigenic C-5 clone (cells) and xenograft C-5 tumor isolated from nude mice.
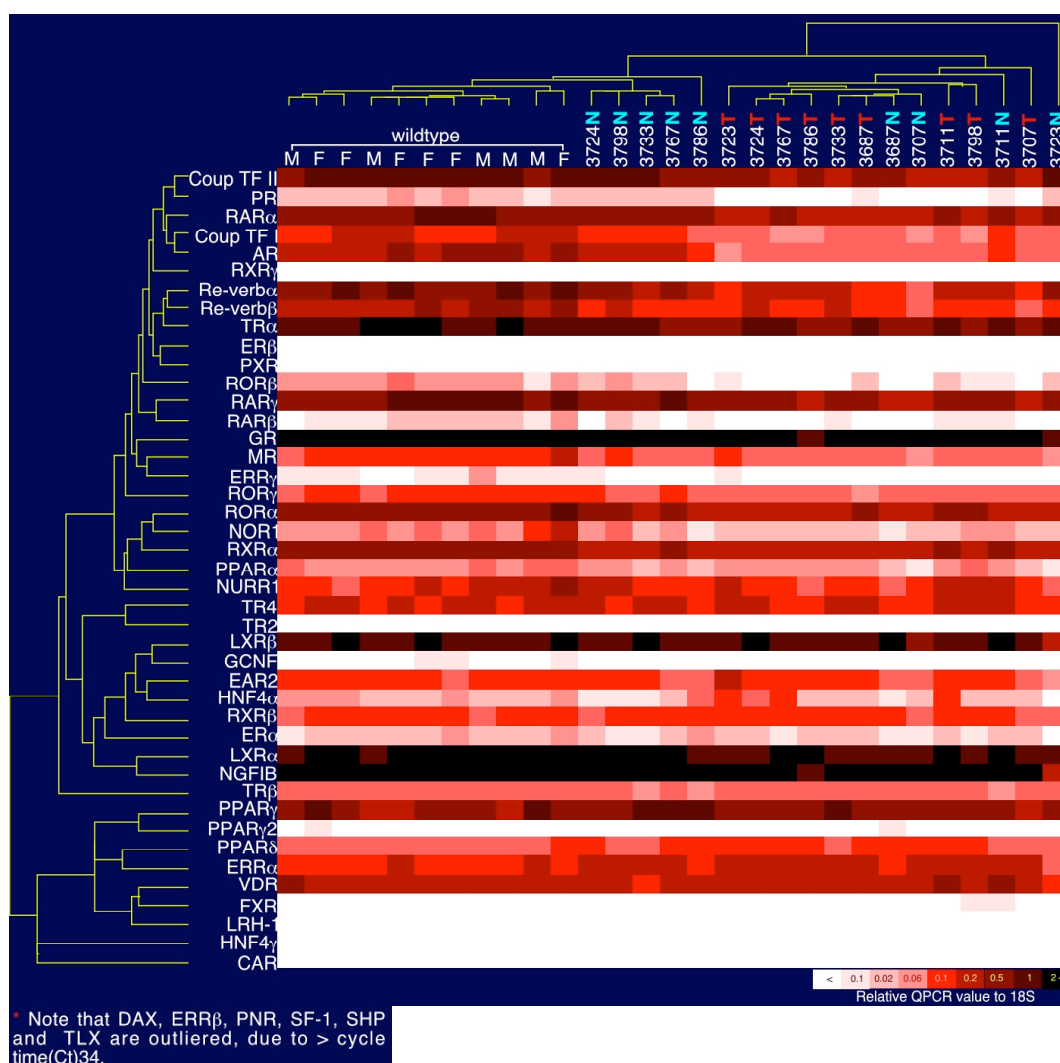
Figure 5.7 Unsupervised Cluster Analysis of NR profile in Mouse Lung Tissues. Expression profiling of the NR superfamily was performed in eight to nine month old mouse lung tissue samples obtained from wildtype and mutant *K-ras*$^{V12}$ transgenic mouse in which lung tissues were separated into normal and corresponding tumor tissues. Hierarchical clustering was applied to both samples and NR expression profile. M: male, F: Female. Each mouse sample has identification number together with N for normal and T for tumor.

A



**Male wt**

*Note that CAR, DAX, ERRβ, HNF4γ, PNR, PXR, SF-1, SHP and TLX are outliered, due to > cycle time(Ct)34.

**K-RAS /Male Normal**

**K-RAS/Male tumor**

**Hierarchical Clustering (Pearson Correlations) of Nuclear Receptors according to Disease Progression**
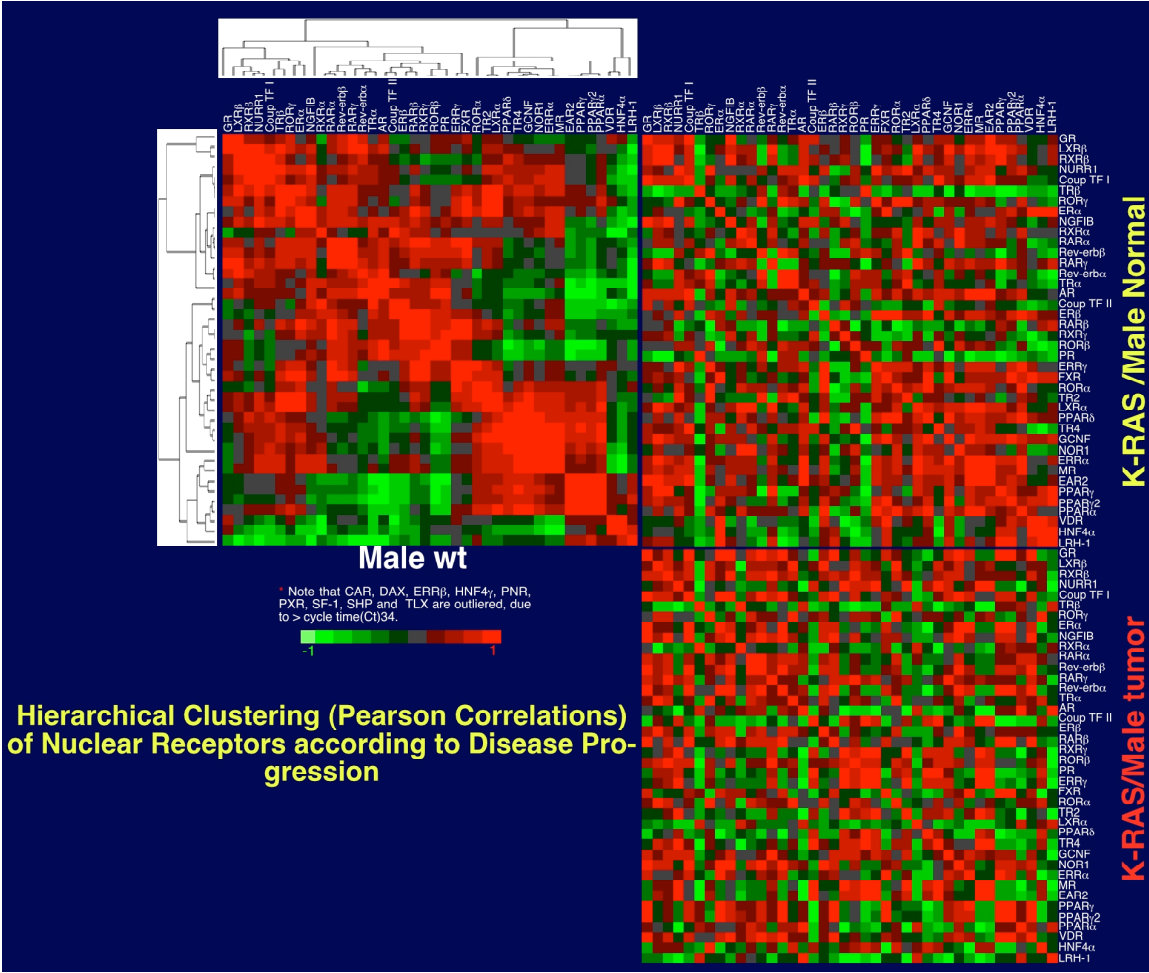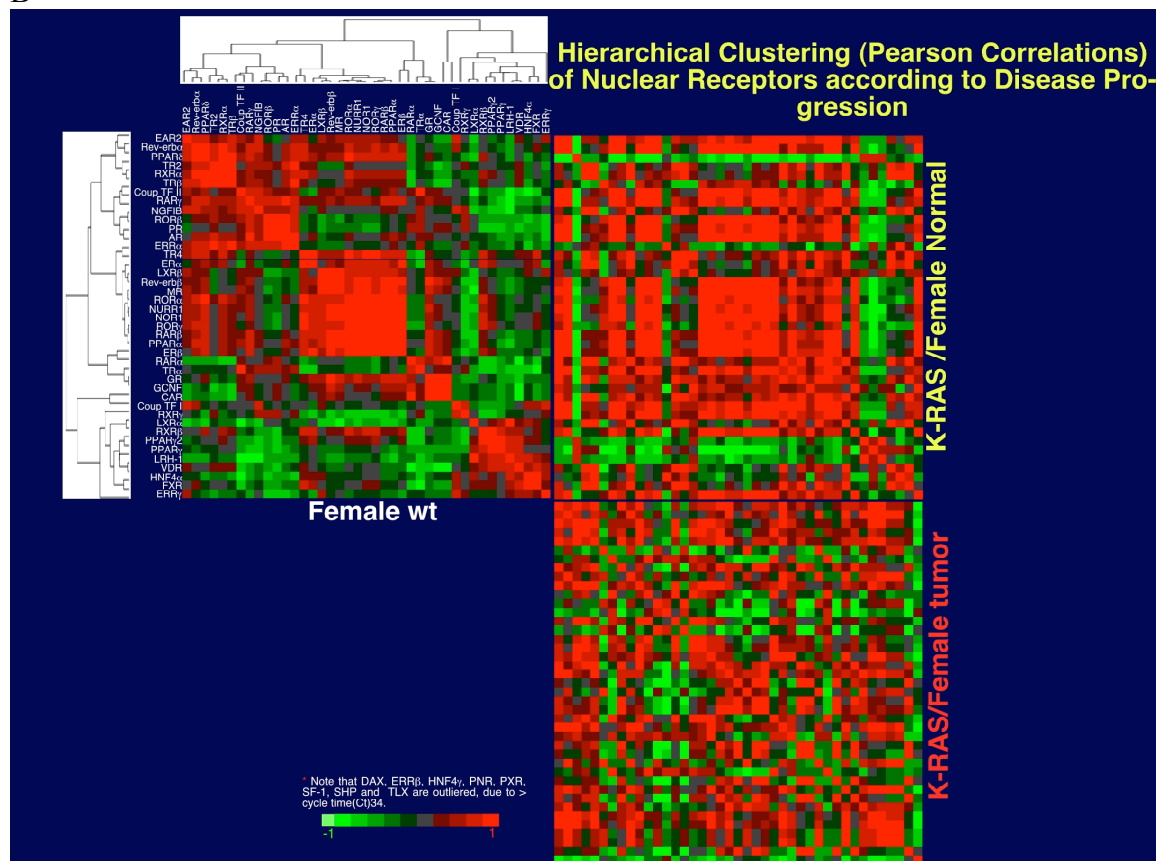
Figure 5.8 Expression Signatures of NR with Disease Progression.
Pearson based correlation coefficients between receptor pairs (vertical vs. horizontal axis) were calculated and applied to hierarchical cluster analysis in wildtype group (A, male; B, female). Both axes showing clustered receptor order in the wildtype group were applied to other tissue samples (i.e., K-ras/normal tissue and K-ras/tumor tissue) with color-code.

# CHAPTER SIX
# PERSPECTIVES AND FUTURE DIRECTIONS

6.1 Introduction

Here I proposed a novel, promising translational approach of utilizing expression profiling of the NR superfamily in lung cancer. Recently, anatomical and temporal profiling of the NR superfamily in mice gave insight to the existence of a higher-ordered mega-network of NRs where finely-tuned balances may play a critical role in maintaining body homeostasis. Thus, as an obvious progression from the previous studies, the current study provides evidence for the NRs that are coordinately involved in these fatal diseases, as well as how NRs expression levels change during disease progression. Through large-scale screening of the NR superfamily, I believe that a new paradigm, i.e., quantitation of specific gene families, has emerged with direct translational applications in clinic. These studies included the well-coordinated organization of various data sources, i.e. qRT-PCR for high-throughput expression analysis, biostatistics and bioinformatics analyses, and a statistically meaningful number of samples. Both biostatistics and bioinformatics tools derived secondary NR datasets such as expression patterns correlated to diagnosis or prognosis, from the primary profile data. After selecting receptors based on expression, ligands were used to

128

evaluate both *in vitro* and *in vivo* potential power in molecular-targeted therapy as well as individualized treatment. Thus, following further preclinical studies using xenograft models and the mouse lung cancer model, clinical trials will be planned to treat lung cancer patients in a variety of ways, single NR targeted therapy, combined NR targeted therapies, and combined treatment with other targeted drugs in clinic. In parallel, preclinical chemoprevention studies should be planned with the consideration that many NRs are relevant to dietary metabolism. More importantly, these various approaches are possible in a customized manner by differentiating individual variations in response to drugs. The notion of using the NR superfamily as a target for translational research is now expanding to other types of cancers. Here, I describe some preliminary data and functional assays in a breast cell line panel.

6.2 Breast Cancer

NR expression was profiled in a panel of 35 breast cell lines including four immortalized mammary epithelial cells and thirty-one breast cancer cell lines. The immortalized cells include pre-immortalized (HME50-5), spontaneously immortalized (HME50-5E), and immortalized cells with hTERt and/or CDK4. NR expression profiles of these breast cell lines displayed distinct but unique expression patterns for both receptors and cell lines (Figure 6.1). The quantitative mRNA expression of both estrogen receptor $\alpha$ and progesterone receptor, the most widely used steroid receptors for diagnosis and prognostic prediction of breast cancer in the clinic, for every breast cell line was compared to immunohistochemistry data documented when the patients were diagnosed. Convincingly,

the data are highly correlated to each other and this was further confirmed with the microarray data (Figure 6.2). In addition, the expression data for estrogen receptor $\alpha$ were functionally confirmed by treating multiple cell lines, including ER$\alpha$-negative and positive cells, with 17$\beta$ estradiol for 3 consecutive days. The data show growth response to estrogen treatment in a receptor dependent manner (Figure 6.3). Further comparative analysis using bioinformatics tools is required for more detailed interpretation.

6.3. Future directions and Perspectives

The NR profile in the breast cell panel seems to be highly informative for translational applications in the clinic and potentially may provide additional or alternative therapeutic targets other than ER$\alpha$ and PR. Along with preclinical studies *in vitro* and *in vivo*, the same approach will be taken with NR expression in breast cancer patient tissue samples as in lung cancer, and followed by design of clinical trials to test whether patients whose tumors have certain NR expression profiles will benefit from selected hormonal treatment/manipulation. It is of interest to investigate if NR expression can differentiate tumor types, and what specific groups of NRs are principal component sets for tumor typing. The spontaneously immortalized HME 50-5E showed increased expressions of PPAR$\gamma$ and COX2, which is similarly observed in immortalized bronchial epithelial cells (Figure 6.4 and Figure 5.3), suggesting that the pathogenic process of epithelial cells in both tissue types may involve PPAR$\gamma$ activation in which a detailed mechanism still remains to be determined. Consistent with our observation, it was recently reported that COX2 activation is significantly involved in late stages of immortalization (Crawford et al., 2004). By the same

token, I believe that it is crucial to investigate any single NR or subset of NRs significantly involved in common type of cancers (e.g. adenocarcinoma), originated from various tissues (lung, breast, colon, and prostate cancers), or metastatic cancers to multiple tissues. Given that therapeutic responses to certain targeted drugs vary or certain treated patients become refractory, it will be of particular interest to see if any NR set changes in expression before and after treatment in the same patient. The answer may provide insights on how the treated tumors become refractory or relapse, and thus how to direct the subsequent treatment schemes. On the basis that many targeted therapies include kinase pathways, currently in progress is profiling of the kinase family (termed 'kinome') for further correlations to the NR superfamily. Using both datasets, combined therapeutic approaches can possibly be planned in preclinical studies followed by clinical applications.

In terms of the notion of cancer stem cells, in spite of controversy for its definition, and given that many NRs are involved in cellular differentiation, it is of interest to see if NRs are relevant to cancer stem cell physiology. Various known cancer stem cells or isolated stem cells from different tissues will be precious sources in which profiling the NR superfamily and further preclinical, biological research should be followed for the use of NRs as clinical targets.

Furthermore, the connection to the microRNA (miRNA) field that current cancer research has been highlighting might be considered. In addition to NRs being involved in transcriptional regulation, thinking about the potential involvement of miRNAs controling posttranscriptional regulation of gene expression is reasonable for understanding higher-level cross-regulation. Using biostatistics analysis, surveying the correlation of individual NRs to

miRNAs in various normal tissues and further comparison in the same tissues with cancer will provide functionally relevant groups of miRNAs and NRs which are potentially cross-targets of miRNAs for NRs or vice versa. Functional analysis based on molecular biology will lead to novel mechanisms of gene regulation involving both repression through miRNAs and direct activation by NR. Recalling the potential mega-network within the NR superfamily for integrated physiological outcomes, further bioinformatics approaches will lead to mechanisms of highly complicated, but finely regulated communication in gene regulation, termed 'transcriptional physiology'. Although the NR superfamily plays central roles in many physiologic processes, further studies to profile expression of generic transcription factors as well as coregulators is extremely important to dissect physiological responses at the transcriptional level. The reconstitution of the expression profile using computational systems could be a valuable systems biology tool for virtual prediction of physiological responses according to stimulus even before preclinical studies. It is necessary to develop such systems for biologic approaches for simulation of the signaling atlas, which could be applied to pharmacologic approaches in biomedical research.

Overall, the work presented here is a blueprint for the above perspectives to be applied and a cornerstone toward understanding the higher level of integrative biology combining transcriptional physiology and systems biology.

A



B
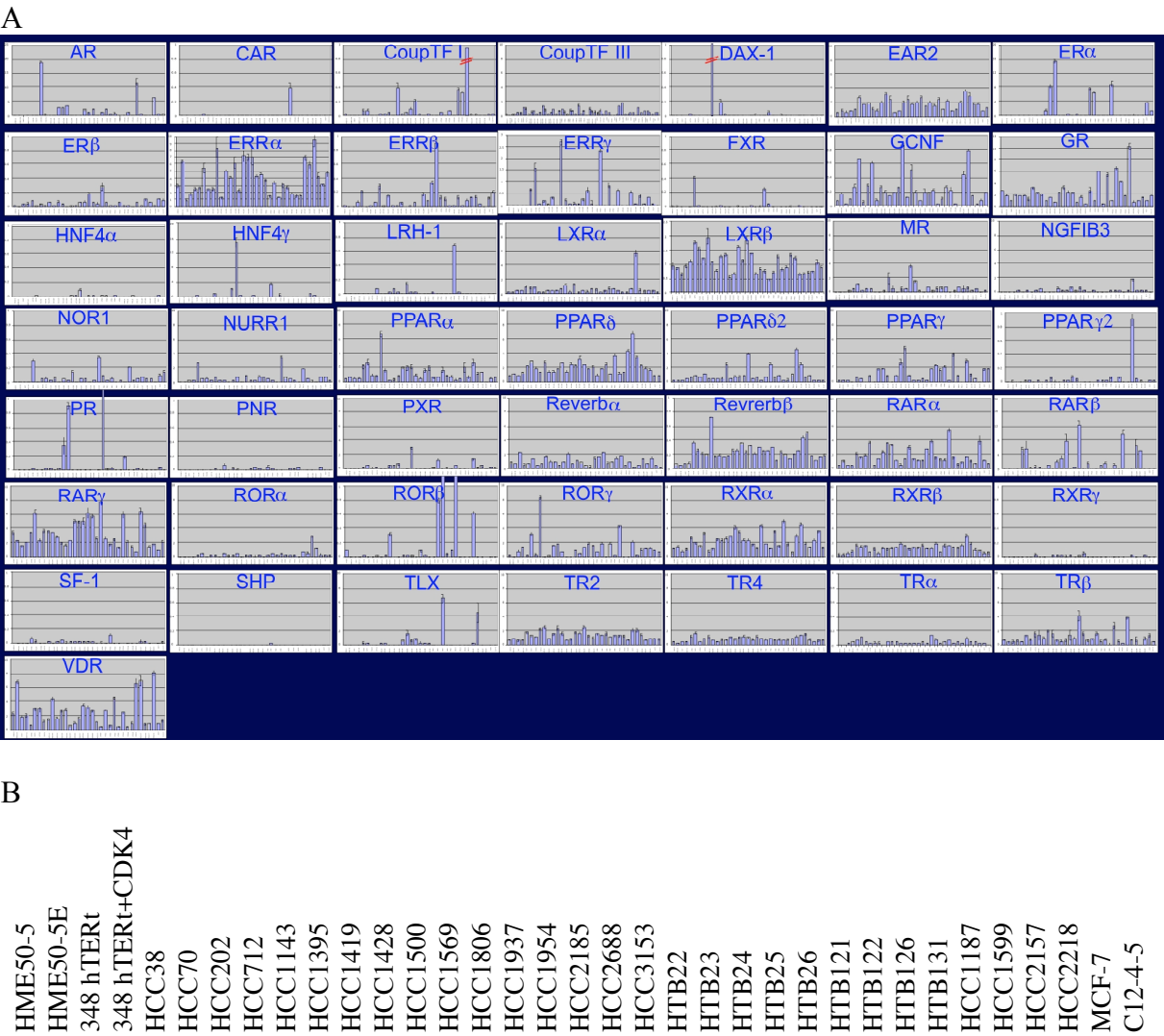


Figure 6.1 Expression Profile of the NR Superfamily in Breast Cancer Panel.
(A) Included in a panel of 35 breast cell lines are four immortalized human mammary epithelial cells and thirty-one breast cancer cell lines. Expression profile of fifty NRs including two isoforms, PPARδ2 and PPARγ2, was surveyed using TaqMan-based quantitative real-time PCR assay. Y-axis represents NR expression relative to18S rRNA. (B) Cell lines are aligned in order as shown on x-axis in figure (A). First four cell lines are immortalized and the rest are breast cancer cell lines.

A



B
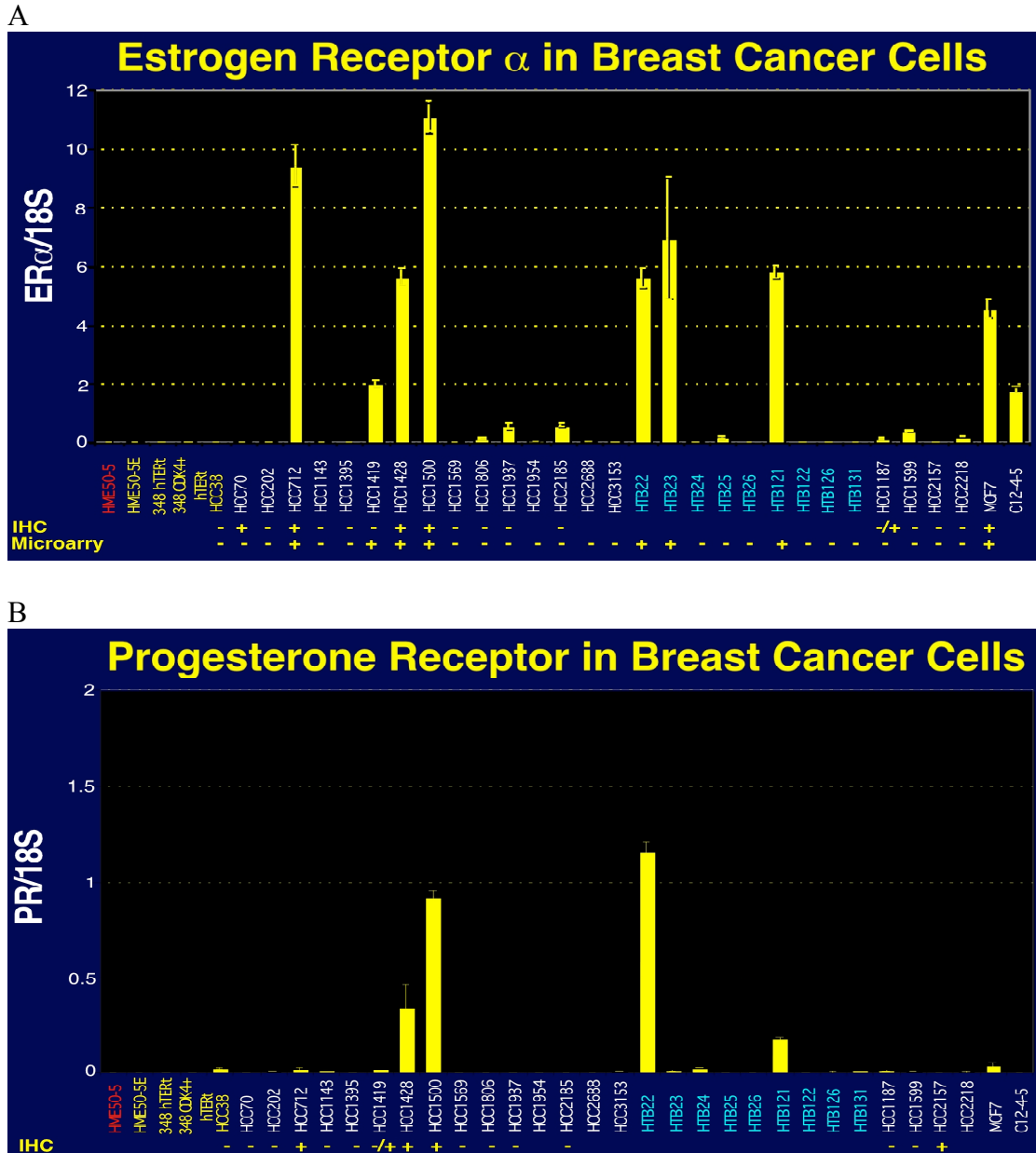


Figure 6.2 Expression Correlations of ERα and PR QPCR Expression with Immunohistochemistery and Affymetrix Microarray Data.
QRT-PCR mRNA expression for both estrogen receptor α (A) and progesterone receptor (B) is compared with Affymetrix Microarray and /or Immunohistochemistry. Y-axis represents relative mRNA expression of nuclear receptor to 18S rRNA.
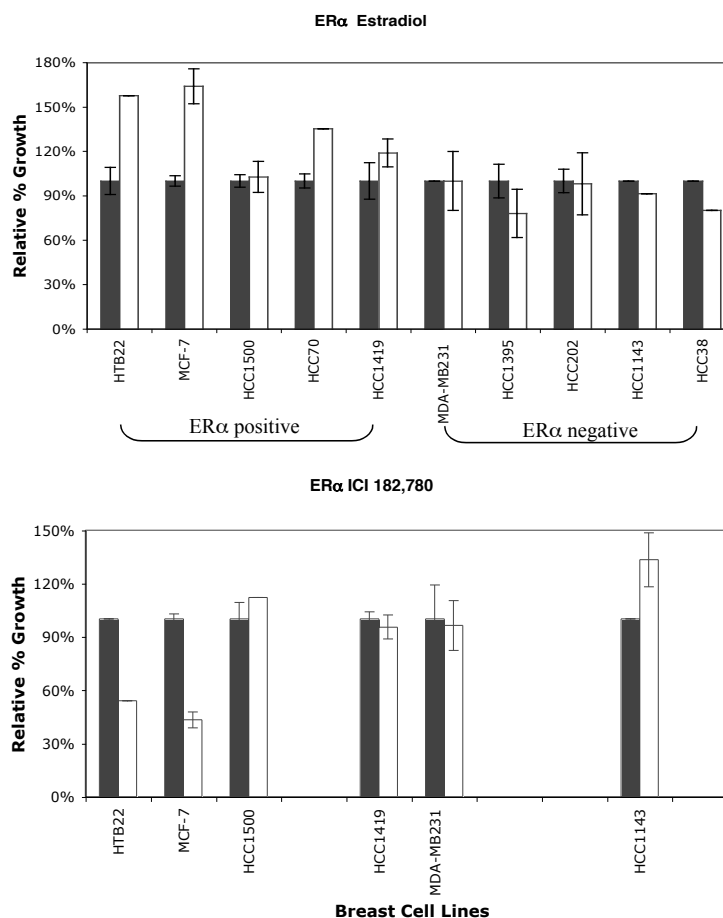
Figure 6.3 Functional Evaluation of ERα Expression.

A subset of breast cell lines expressing or not expressing ERα were treated with ligands and assayed for cell growth response. Final concentration of 10nM for both ligands (17β-estradiol as an agonist and ICI 182,780 as an antagonist) was used for three consecutive days before cell counting with trypan blue exclusion. All relative % growths were represented by normalizing treated cells with ligands by same cells with vehicle treatment.
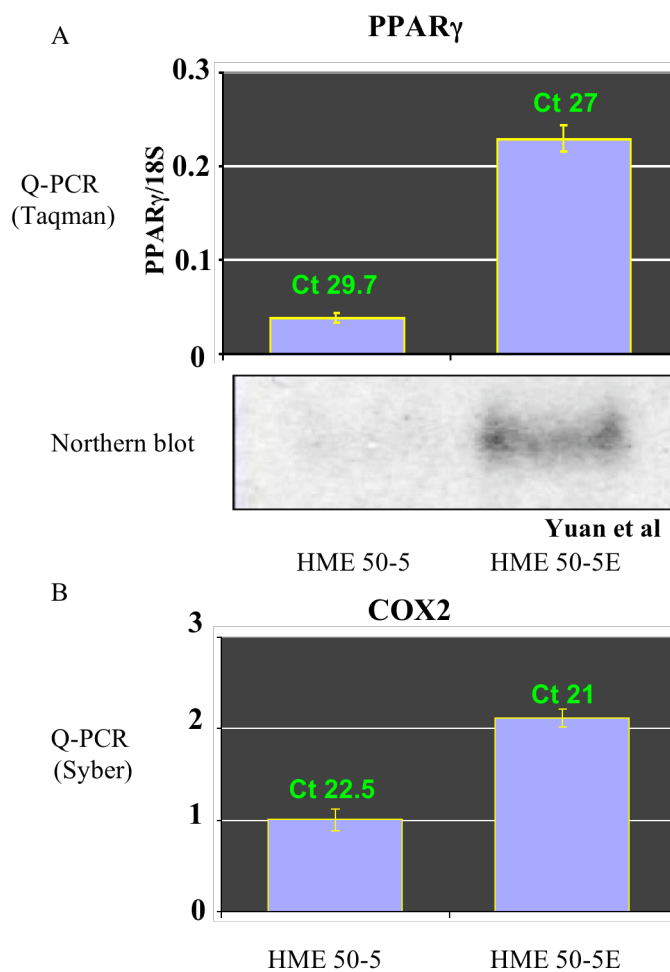
Figure 6.4 Expression of PPARγ and COX2 in HME 50-5 and HME 50-5E. Quantitative mRNA expression for PPARγ (A) and COX2(B) was examined in HME 50-5 and HME 50-5E cells. Northern blot assay shows mRNA expression of PPARγ in (A). Ct: cycle time.

# CHAPTER SEVEN

# Materials and Methods

## 7.1 Materials

### 7.1.1 Lung Panel

Primary human bronchial epithelial cells (HBECs), immortalized HBECs, and tumorigenic clones are described in section 2.2.2 and schematized in figure 2.3. Note that successive clonogenic analysis established more immortalized but non-tumorigenic clones which are believed to be intermediate stages toward tumorigenic clones such as C1 and C5 HBECs. Our lung cancer cell panel included five primary bronchial epithelial cell lines (under passage five) including normal bronchial epithelial cells (NHBEC) purchased from Clonetics, ten immortalized HBEC lines with CDK4 and hTERt in the presence or absence of various oncogenic alterations, and forty-one lung cancer cell lines. The lung cancer cell lines consist of fourteen SCLCs and twenty-seven NSCLCs, for which their identities were confirmed by DNA fingerprinting. All RNAs were extracted using Qiagen RNeasy Midi kit (Qiagen Sciences, MD), given unique identification codes, and used for both microarray and qRT-PCR experiments.

7.1.2 Patient Tissue Sample

Thirty primary tumors and corresponding normal tissues, comprising 23 adenocarcinomas and 7 squamous cell carcinomas, were obtained and further processed for RNA under the approval of the Internal Review Board (IRB) at MD Anderson Cancer Center and IRB at UT Southwestern Medical Center, together with written informed consent for each subject. Sixteen patients were diagnosed with stage I disease, five patients with stage II disease, five patients with stage III disease, and four patients with stage IV disease. All tissues were stored at −80 °C after being snap frozen. Cancerous regions were microdissected and RNAs were isolated using the Qiagen RNeasy Mini Kit (Qiagen Sciences, MD).

7.1.3 Mouse Lung Cancer Model

To better understand the pathogenic relevance of the NR superfamily in an *in vivo* mouse model with constitutive expression of oncogenic *K-ras$^{V12}$* that spontaneously develops adenoma and adenocarcinoma at later stages, several groups of lung tissue samples from the mice fed the same diet were collected from both genders from early age (4 to 6 weeks), middle age (4 to 6 months), and old age (8 to 10 months) animals that had developed mostly adenocarcinoma. Each group consisted of five wild types and five mutants in both genders. All mutant samples were pair-matched and RNAs were prepared and analyzed for RNA quality control using the Agilent 2100 bioanalyzer (Quantum Analytics Inc., Foster city, California).

7.2 Methods

Cell cultures, western blot assay and RNA extraction

NHBEC, HBECs UI, and BEAS2B cells (American Type Culture Collection, Manassas, Virginia) were cultured in BGEM® medium (Cambrex Bio Science, Walkersville, Maryland) supplemented with the supplied nutrients according to the manufacturer's instructions. HBECs were grown in keratinocyte-SFM (Invitrogen Corp., Grand Island, New York) with the supplements (e.g. EGF, pituary gland extract) supplied. Lung cancer cells were maintained in RPMI medium (Invitrogen corp., Grand Island, New York) medium with 5% or 10% heat-inactivated fetal bovine serum (ΔFBS) (Gemini Bio Products, Woodland, California). Note that the FBS in the culture continuously used contained 17β-estradiol and testosterone which were in the range of physiological concentration (0.1 ~ 1 nM). All genomic DNAs were fingerprinted to confirm the identity of each of the cell lines used in this study. Following standard protocols for western blot assay, a total of 50 μg of whole cell lysate was assayed with mouse monoclonal anti-PPARγ1 (Tanaka et al., 2002) to confirm PPARγ protein expression. For cyclin D1 detection, a total of 50 μg of whole cell lysate from H1770 and H2347 was assayed with mouse anti-cyclin D1 (Santa Cruz, sc-8396) antibody at three different time points: 0hr, 24hr and 48hr after 3 μM troglitazone treatment. Total RNAs were purified using Qiagen RNeasy Midi kit (Qiagen Sciences, MD) following the manufacturer's instructions. The quality and concentration of each RNA was assessed using either ethidium bromide-stained agarose gels or the Agilent 2100 bioanalyzer (Quantum Analytics Inc., Foster city, California).

**Reverse Transcription and Quantitative Real-time PCR assay**

All cDNAs were prepared for quantitative real-time PCR (TaqMan® method) as described (Bookout et al., 2006). Briefly, 2 μg of total RNA was DNAse-treated with 2 U of DNAse I in a final volume of 20 μl containing 4.2 μM $MgCl_2$ . The reverse transcription reaction was performed in a 100 μL final volume, followed by the addition of 100 μL of DEPC-H2O. Human universal cDNA or tissue specific cDNA (e.g. retina for PNR and liver for SHP) was used to construct a standard curve of the following concentrations, no template control, 0.008, 0.04, 0.2, 1, 5, 25 ng for 18S and NTC, 0.016, 0.08, 0.4, 2.0, 10, 50 ng for each NR. These quantities are based on the RNA concentrations used for the RT reaction. A negative RT sample, a control for genomic DNA contamination, was also included for both 18S and NR. For each sample, 10 ng of cDNA was assayed in triplicate wells of a 384-well plate. The final forward and reverse primer concentrations used were 75 nM for 18S rRNA and 300 nM for all of the NR primers. Primers for each mRNA were designed using Primer Express Software and were validated as previously described (Table 2.1 and http://www.NURSA.org).

**QPCR Data analysis**

The data was imported into Microsoft Excel® and, first, calculated for PCR efficiency (*e*), $e = 10^{[-1/slope]}$ where the slope was obtained from the standard curve calculated by the SDS instrument for the endogenous 18S reference and target NR. Secondly, relative quantity was calculated by *quantity* = $(e)^{-Ct}$. The calculated quantities were averaged (*avg*), and the

standard deviations (*stdev*) and coefficients of variation (*CV=stdev/avg*) were determined for the 18S and NR of each sample. Points which showed > 17% CV were considered outliers and removed. Then the normalized value for each NR expression was calculated using *normalized value = NR quantity avg / 18s quantity avg*. The standard deviation of the normalized value (S.D.) was calculated as S.D. = (normalized value) X $\{($CV of reference$)^2 + ($CV of NR$)^2\}^{1/2}$. The normalized values are represented as a bar graph.

**Microarray data analysis**

To investigate the signatures of potential target gene expression in the context of steroid receptors (e.g., AR, ERα) and PPARγ expression, microarray assays were performed using Affymetrix U133AB chips for 48 lung cell lines. Expression data were analyzed using unsupervised clustering method. Lung cancer cell lines employed for the correlation of receptors selected are as follows: a total of 12 lung cancer cell lines for AR, four positive cell lines (H2122, H1993, H460, H1184) and 8 negative cell lines (H2009, HCC827, HCC1195, H1607, H1299, H2882, HCC366, H289); a total of 28 cell lines for ERα, 5 positive cell lines (H1607, H1993, HCC78, H2052, HCC1195) and 23 negative cell lines (H1299, H157, H2882, H1819, H2087, H358, HCC44, H2887, HCC15, HCC366, HCC461, H146, H1672, H2107, H889, H289, H187, H82, H1963, H1184, H2171, HCC970, H2227); a total of 10 cell lines for PPARγ, 5 high-expressing cell lines (H1993, H2347, Calu-1, H2887, H2882) and 5 low-expressing cell lines (HCC1195, H1770, H2227, H187, H1299). Matrix 1.29 software (developed by Dr. Luc Girard) or Eisen lab Tree software (http://rana.lbl.gov/EisenSoftware.htm) was used for unsupervised clustering analysis of

QPCR data of the NR expression and Affymetrix microarray data. In addition, Matrix 1.29 was utilized to pull out groups of genes showing more than 2-fold differences between receptor positive and negative cell lines. Pearson correlations of each gene to each receptor expression level was calculated and followed by significance test for the correlation: $P$=TDIST(ABS(r/SQRT((1-r^2)/(n-2))),n-2,2) where r is Pearson correlation value and n is sample number tested. Genes successfully filtered-through (P < 0.05) were highlighted in yellow.

**Cell growth assay**

To assess the growth response to ligand treatments, cell lines MCF-7, HCC78, and H2052 which express ERα and H1299, H2009 which show no ERα, were selected for ERα evaluation and maintained in phenol-red free RPMI medium containing 5% heat-inactivated, charcoal-stripped fetal bovine serum (ΔFBS). For the assay, 1 X $10^5$ cells were split into 6-well plates and grown in phenol-red free RPMI media containing 5% charcoal-stripped, heat-inactivated ΔFBS. Agonist, 17β-estradiol, (Sigma, St. Louis and Missouri) or antagonist, ICI 182, 780, (TOCRIS, Ellisville, Missouri) was added every day for 3 days in 4 doses (0, 0.1nM, 10nM, 1μM) in triplicate. Relative % of cell growth response was calculated by counting Trypan blue-excluding cells (Sigma, St. Louis, Missouri). For AR evaluation, all assays were performed in the same way as ERα except cells were maintained in phenol-red containing media. Agonist, dihydrotestosterone (DHT), (Sigma, St. Louis, Missouri) was added using the same concentrations as ERα ligands except for the cells LnCaP and H2009 (0, 0.01 nM, 1 nM, 100 nM). To assess the functional and preclinical implications of PPARγ

expression, the MTT (Sigma-Aldrich, St. Louis, Missouri) assay was employed to measure cell proliferation. A number of cells which are low- (H1770, HCC1195, H1299) or high-expressers (H1993, H2347, Calu-1) were split into 96-well plates containing phenol red-free RPMI media containing 5% charcoal-stripped $\Delta$FBS in final volume of 100 μl per well. Cells were treated with troglitazone at the concentrations 0, 0.1, 0.3, 1, 3, 10 μM or 0, 0.1, 1, 10 for Calu-1 cells in a final volume of 125 μl. MTT assays were performed by measuring absorbance of MTT metabolites at 560 nm. Relative % growth was calculated for each dose versus vehicle treatment.

**In vivo xenograft experiment**

Athymic nude mice (5–6-week-old females) were purchased from Charles River Laboratories (Wilmington, MA) and maintained in sterile conditions. The care and treatment of experimental animals were in accordance with institutional guidelines. H1770 and H1299 (with no expression of PPARγ), and H2347 (with high expression of PPARγ) ($2 \times 10^6$) cells were subcutaneously injected into the right flank areas of mice on Day 0 (Nakagawa et al.)

. The mice were randomly divided into two groups (n=4) and treated with intraperitoneal administration of 25 mg/kg pioglitazone (Actos, Takeda Pharmaceutical Company, Osaka, Japan) dissolved in 10% DMSO or the same volume of DMSO vehicle control four times a week. Tumor volumes were directly measured twice a week from Day 8 with calipers and calculated by the formula $\pi/6$ X (large diameter) X (small diameter)$^2$. The mice treated with vehicle were sacrificed at day 37 before tumor necrosis occurs. The

difference in tumor volume between different treatment groups was statistically analyzed

using the student $t$-test in SPSS 11.5 software (SPSS Inc., Chicago, IL).

Table 7.1.Quantitative real-time PCR primer and probe sets of 50 human NRs.

Reference sequence accession numbers are included.

| Receptor | FormalmRNA Name | Accession # | QPCR Primers | TaqMan Probe |
|---|---|---|---|---|
| 18S | | X00686 | 5'accgcagctaggaataatgga3'<br>5'gcctcagttccgaaaacca3' | VIC-accgcggttctattt |
| AR | NR3C4 | NM_000044 | (3525)5'gaattcctgtgcatgaaagca3'<br>(3610)5'cgaagttgatgaaagaatttttgatt3' | 6FAM-ctactcttcagcattattcca |
| CAR | NR1I3 | NM_005122 | (771)5'ttcatggtactgcaagtcatcaa3'<br>(856)5'ttgagaagggagatctggtcttc3' | 6FAM-ttactaaggacctgcccgtc |
| COUP-TFα | NR2F1 | NM_005654 | (1053)5'acagctgcctcaaagccatc3'<br>(1167)5'tcacgtactcctccagtgcg3' | 6FAM-cacatcgagagcctgcag |
| COUP-TFβ | NR2F2 | NM_021005 | (1249)5'aaggcgctgcacgttgac3'<br>(1349)5'ctttccacatgggctacatcag3' | 6FAM-tagtcctgtttcacctcagat |
| DAX | NR0B1 | NM_000475 | (1068)5'ccaggtccaagccatcaagt3'<br>(1175)5'ggcacgtccgggttaaaga3' | 6FAM-aggagtacgcctacctc |
| COUP-TFγ | NR2F6 | NM_005234 | (352)5'agggctgcaagagctttttc3'<br>(427)5'tggcagtcacggttgga3' | 6FAM-caacctcagctacacctg |
| ERα | NR3A1 | NM_000125 | (989)5'agagaagtattcaaggacataacgactatat3' (1066) 5'tcttcctcctgttttatcaatgg3' | 6FAM-cagccaccaaccagt |
| ERβ | NR3A2 | NM_001437 | (1316)5'aagttggccgacaaggagtt3'<br>(1391)5'acaggctgagctccacaaag3' | 6FAM-ccaagaagattccc |
| ERRα | NR3B1 | NM_004451 | (1143)5'gcgagaggagtatgttctactaaagg3'<br>(1218)5'agcctcggcatcttcgat3' | 6FAM-ccaattcagactctgtgcac |
| ERRβ | NR3B2 | NM_004452 | (719)5'gaggactatccaagggaacattg3'<br>(839)5'catccccactttgaggcatt3' | 6FAM-cgagatcaccaaacgg |
| ERRγ | NR3B3 | NM_001438 | (182)5'gctaacactgtcgcagtttgaa3'<br>(357)5'cgaacagctggaatcaatgtg3' | 6FAM-tctgcagaatgtcaaaca |
| FXR | NR1H4 | NM_005123 | (897)5'tgtcgactaaggaaatgcaaaga3'<br>(997)5'tgctgcttcacattttttctca3' | 6FAM-cttgttaactgaaattca |
| GCNF | NR6A1 | NM_033334<br>NM_001489<br>NM_001489<br>NM_033335 | (1279)5'atcgagcggctcatctacct3'<br>(1379)5'atatcttgatttaggaagttaattgctttc3' | 6FAM-agttccatcagctaaag |
| GR | NR3C1 | NM_000176 | (1280)5'tccctggtcgaacagtttttt3'<br>(1357)5'agctggatggaggagagctt3' | 6FAM-ccagcatgagaccagat |
| HNF4α | NR2A1 | NM_000457 | (536)5'tgcaggctcaagaaatgctt3'<br>(591)5'tcattctggacggcttcctt3' | 6FAM-catgccagcccgga |
| HNF4γ | NR2A2 | NM_004133 | (360)5'tgggtgcaagggtttcttc3'<br>(426)5'ccgactgaacctgcaagaata3' | 6FAM-cagcattcgtaagagtcaca |
| LRH-1 | NR5A2 | NM_003822 | (469)5'cagagaaagcgttgtccttactg3'<br>(569)5'ttattccttcctccacgcatt3' | 6FAM-tcggcccttacagcttct |
| LXRα | NR1H3 | NM_005693 | 5'cccttcagaacccacagagatc3'<br>5'gctcgttccccagcatttt3' | 6FAM-ccacaaaagcggaaaa |
| LXRβ | NR1H2 | NM_007121 | (1099)5'cgctaagcaagtgcctggtt3'<br>(1199)5'gcctggctgtctctagcagc3' | 6FAM-ccctcctgaaggcat |
| MR | NR3C2 | NM_000901 | (2775)5'ccaaatcagccttcagttcgt3'<br>(2876)5'ttgaggccatcctttggaat3' | 6FAM-ttagtagcagcaaaactttca |
| NGFIB3 | NR4A1 | NM_173158 | (139)5'cccttcgtgcggttgtct3'<br>(220)5'ggcttggatcacgggcatct3' | 6FAM-cctttttccagggtcaaa |

| Receptor | FormalmRNA Name | Accession # | QPCR Primers | TaqMan Probe |
|---|---|---|---|---|
| NOR1 | NR4A3 | NM_006981 NM_173198 NM_173199 NM_173200 | (1741)5'aagagacgtcgaaaccgatgt3' (1833)5'tttcagactatctgtacggacaacttc3' | 6FAM-taccattccaacactgag |
| NURR1 | NR4A2 | NM_006186 | (1625)5'gcccaaagccgaccaa3' (1700)5'ggacctgtatgctaatcgaagga3' | 6FAM-ctgcttttgaatcagct |
| PNR | NR2E3 | NM_016346 NM_014249 | (456)5'tgtgccccgtggacaa3' (542)5'cggcgtcctggttcatc3' | 6FAM-acttcttcagccgg |
| PPARα | NR1C1 | NM_005036 | (148)5'acgtgcttcctgcttcatagat3' (218)5'caccatcgcgaccagatg3' | 6FAM-agctcggcgcacaa |
| PPARδ | NR1C2 | NM_006238 NM_177435 | 5'cagtactgccgcttccagaa3' 5'ggccatccgaccaaaacg3' | 6FAM-catgtcacacaacgctat |
| PPARδ2 | | NM_177435 | (128)5'ccaacagatgaagacagatgca3' (197)5'ctgaacgcagatggacctcta3' | 6FAM-tgatgggaaccaccc |
| PPARγ | NR1C3 | NM_005037 NM_015869 | 5'agatccagtggttgcagattaca3' 5'ggagatgcaggctccacttt3' | 6FAM-tcaagagtaccaaagtgca |
| PPARγ2 | | NM_015869 | (163)5'tgacccagaaagcgattcct3' (263)5'caaagttggtgggccagaa3' | 6FAM-tatagcacaccatcccca |
| PR | NR3C3 | NM_000926 | (1927)5'tgggagctgtaaggtcttctttaa3' (2007)5'acgatgcagtcatttcttcca3' | 6FAM-aagggcagcacaact |
| PXR | NR1I2 | NM_003889 | (2984)5'cccagcctgctcataggttc3' (3064)5'gggtgtgctgagcattgatg3' | 6FAM-catggctatgctcaccgag |
| RARα | NR1B1 | NM_000964 | 5'cagcaccagcttccagttagtg3' 5'ctgctgctctgggtctcaatg3' | 6FAM-tatagcacaccatcccca |
| RARβ | NR1B2 | NM_000965 NM_016152 | (1390)5'cagctcctgcctttggaaa3' (1489)5'cttttgtcggttcctcaaggt3' | 6FAM-ttaatctgtggagaccgcca |
| RARγ | NR1B3 | NM_000966 | 5'ggaacaagaagaagaaagaggtgaa3' 5'ttggtgatgagctcttctaactga3' | 6FAM-cctgacagctatgagc |
| Rev-erbα | NR1D1 | NM_021724 | (2025)5'tgaccaagtcaccctgcttaag3' (2083)5'aagcaaagcgcaccatca3' | 6FAM-cacctcaaaggtgccag |
| Rev-erbβ | NR1D2 | NM_005126 | (773)5'tcagcaatgtcacttcaaaa3' (835)5'ccaaaccgaacagcatctctt3' | 6FAM-acattccaacagacagaca |
| RORα | NR1F1 | NM_134261 | 5'ctgactgaagatgaaattgcattatt3' 5'gcagccatgagcgatctg3' | 6FAM-ctgcatttgtactgatgtc |
| RORβ | NR1F2 | NM_006914 | (1186)5'ttcaacaatgggcagttagca3' (1266)5'ccaaatgggacttaatgatgttct3' | 6FAM-ccatgactgaaatcgaccga |
| RORγ | NR1F3 | NM_005060 | (951)5'gcagcgctccaacatcttc3' (1031)5'gcacaccgttcccacatct3' | 6FAM-cacatggacttcctctggtag |
| RXRα | NR2B1 | NM_002957 | (795)gagcccaagaccgagaccta3' (895)agctgtttgtcggctgctt3' | 6FAM-ccagctcgccgaac |
| RXRβ | NR2B2 | NM_021976 | (654)5'agcccccagattaactcaaca3' (800)5'gattgcacatagccgtttgc3' | 6FAM-atgtgaagccaccagtct |
| RXRγ | NR2B3 | NM_006917 | (215)5'gaagtttcccgcaggctatg3' (278)5'tgatgggctcatggatgtaga3' | 6FAM-aggctcccctggcc |
| SF-1 | NR5A1 | NM_004959 | (386)5'ttctgccgcttccagaaat3' (486)5'ttgtacatcggcccaaacttt3' | 6FAM-aagccgtgcgcgct |
| SHP | NR0B2 | NM_021969 | (529)5'cctgcctgaaagggaccat3' (605)5'ctgcaggtgcccaatgtg3' | 6FAM-ccaggcctccaagc |
| TLX | NR2E1 | NM_003269 | 5'tgatgctaacactctactggctgta3' 5'cagcttctgggaatctgtgttg3' | 6FAM-cacgttcatgccag |

| Receptor | Name | FormalmRNA Accession # | QPCR Primers | TaqMan Probe |
|---|---|---|---|---|
| TR2 | NR2C1 | NM_003297 | (998)5'gcagaccaacggtgatgtttt3'<br>(1057)5'ccaggattcaatgcttttgc3' | 6FAM-aagggcatttgacactc |
| TR4 | NR2C2 | NM_003298 | (816)5'gatgggcatgaaaatggaatc3'<br>(881)5'ggtttctcccgttgcacat3' | 6FAM-cgttcactctgcacag |
| TRα | NR1A1 | NM_003250 | (1372)5'cacggaagtggctctgctg3'<br>(1472)5'gcaggtacgcctcctgactc3' | 6FAM-taatgtcaacagaccgctc |
| TRβ | NR1A2 | NM_000461 | (201)5'tgcgtgggtgccaagt3'<br>(264)5'cctttttttcactgacatctccttct3' | 6FAM-ccacacatgatttaatgaat |
| VDR | NR1I1 | NM_000376 | (547)5'ccccacctactccgacttct3'<br>(605)5'ctccaccatcattcacacgaa3' | 6FAM-ccagttccggcctc |

# CHAPTER EIGHT
## Biostatistics and Bioinformatics

**8.1 Introduction**

Biostatistics and bioinformatics have become necessary tools to analyze large amount of datasets generated by high-throughput technologies such as qRT-PCR and microarray experiments. To better understand the high level of complexity of the transcriptional control of nuclear receptors or the correlation between nuclear receptors, several statistical tools were employed in my research. Although a variety of biostatistics and bioinformatics tools are currently available, here I only focus on certain tools used in this thesis.

8.2. Biostatistics

8.2.1 Pearson Correlations

The Pearson correlation (r), called a linear or product-moment correlation, is defined as 'the sum of the products of the standard scores of the two measures divided by the degrees of freedom' (Encyclopedia). The standard score (z) is expressed as $z = (x-\mu)/\sigma$, with x for individual score, $\mu$ for population mean, and $\sigma$ for standard deviation. The degree of freedom is n-1, substracting 1 from total number of events (n). Thus, the Pearson correlation

coefficient can be rewritten $r_{xy} = \{\Sigma(x_i-\bar{x})(y_i-\bar{y})\}/(n-1)S_xS_y$, where $\bar{x}$ and $\bar{y}$ are means of

sample $x_i$ and $y_i$, $S_x$ and $S_y$ are standard deviations of each sample. Basically this equation

considers the mean value of each sample and thus termed 'centered Pearson Correlation',

whereas the term 'uncentered' is used for cases not including a sample mean value. The

equation mathematically determines the extent to which values of the two variables are

'proportional' to each other. It ranges from –1 to +1 which represents perfect negative and

positive linear relationship, respectively, between two sets of variables. Statistical

significance of the correlation coefficient is determined using $P$=TDIST(ABS(r/SQRT((1-

r^2)/(n-2))),n-2,2) where r is a Pearson correlation value and n is sample number tested. The

squared coefficient ($r^2$) is often used for the proportion of common variation in the two

variables (http://www.statsoft.com/textbook/stbasic.html). When there are more than two

variables, the correlations for every pair can be put into a matrix, called a correlation matrix as

in Figure 2.2 and 3.9. While this parametric correlation (Pearson correlation) is based on a

normal distribution of the samples, non-parametric correlation methods such as Spearman's $\rho$

and Kendall's $\tau$ may be considered for a violation of sample normality, which will not be

discussed in detail here. Also, Spearman *R* and Kendall *tau* assumes that the variables were

measured on at least rank order scale. However, Spearman *R* and Kendall *tau* are usually not

identical in magnitude because their underlying logic and computational formulas are

different. More importantly, while Spearman *R* can be thought of as the regular Pearson

product-moment correlation coefficient as computed from ranks, Kendall *tau* rather

represents a *probability*. In addition, whereas the observed data are in the same order for the

two variables in Spearman R, the Kendall *tau* reveals the probability that the observed data are in different orders for the two variables.

8.2.2  Cox Regression model

This semi-parametric survival model has been a popularly recognized technique for examining the dependence of patient survival on various explanatory variables. Due to the benefit of handling multiple covariates or various $3^{rd}$ confounding factors which potentially affect survival rate, basically this model statistically determines the influencing power of the individual explanatory variables on the dependent variables, for example, survival rate or treatment effectiveness. This modeling-based approach is a combined statistics of logistic regression analysis and survival rate, and thus overcomes multiple limitations involving inability of handling continuous covariates and requirement of normal distribution in the parametric log rank test. In addition to the survival relevance adjusted by multiple covariates, the Cox regression model provides a more precise estimate of the treatment effect on survival under considerations of several prognostic variables. In this thesis, the Cox regression model is used to determine the regression coefficient (or hazard ratio) for each explanatory variable on  survival. The mathematical implication of the term 'regression' involves literally reverse-process from the results to the cause, statistically explaining the relationship between the dependent variable and the explanatory variable. If an estimated hazard rate ratio (or regression coefficient) is greater than 1 (HR $\geq$ 1), it indicates that the explanatory variable is associated with the increased risk for the event of interest (e.g., death). Conversely, the hazard ratio less than 1 (HR $\leq$ 1) represents a decreased hazard for the occurrence of the

event, while the ratio of 1 indicates no correlation between the variable and the event. The proportional hazard ratio can be obtained from the ratio of two hazard functions, for example, the ratio of hazard function of tumor tissue to hazard function of normal tissue in the same patient. The hazard function used in this thesis can be rewritten, $h(t) = h_0(t) \cdot \exp(\beta_{age} \cdot Age + \beta_{stage} \cdot stage + \beta_{gender} \cdot gender + \beta_{recurrence} \cdot recurrence + \beta_{smoking} \cdot smoking + \beta_{histology} \cdot histology + \beta_{NR} \cdot NR)$ where the $h_0(t)$ is the baseline hazard function, usually representing the probability of death when all explanatory variables are zero. If compared to the hazard ratio of tumor vs. normal like this thesis, the baseline hazard function disappeared as shown in $HR = h_t(t)/h_n(t) = h_0(t)\exp(\beta \cdot x_{tumor})/h_0(t)\exp(\beta \cdot x_{normal}) = \exp(\beta(x_{tumor} - x_{normal}))$.

The hazard function represents the probability of the event (e.g., death) occurring for the individual patient in a given time interval. Therefore, the probability can be interpreted as risk of dying at time t (called 'hazard ratio'). The underlying assumption of the hazard function is a constant relationship, or proportional, between the explanatory variables and the dependent variable. Most features are shared between multiple regression and Cox regression where is allowed to consider more than one explanatory variable at any given time but not in multiple regression mode. The statistical significance is assessed using the chi-square test which is a non-parametric test for the bivariate tabular analysis.

8.2.3   Principal Component Analysis

From the expression profile of the NR superfamily, finding the expression pattern of a subset of NRs relative to demographic features in a panel of patient samples is important for understanding certain NRs in tumor pathogenesis and thus for further clinical applications. A

principal component analysis (PCA) was performed to identify principal component sets of

NRs highly relevant to patient survival. Here I introduce the basic concept and process of

PCA using the NR expression profile in tissue samples of lung cancer patients. Considering

one nuclear receptor in one dimension in the tissue panel, total calculations of different

covariance values from 50 dimensions reveal $_{50}C_2$ =50!/((48!)*2) =1225 covariance values

because individual covariance represents variations from the mean with respect to each other.

The formula is as follows;

$$Var(X) = \left\{ \sum_{i=1}^{n}(X_i-\overline{X})(X_i-\overline{X}) \right\} / (n-1)$$

$$cov(X,Y) = \left\{ \sum_{i=1}^{n}(X_i-\overline{X})(Y_j-\overline{Y}) \right\} /(n-1)$$

, where X and Y may represent any two NRs and n is 30. Then, all covariances were

reconstituted on matrices, called covariance matrix, using

$C^{m \times n} = (C_{i,j}, C_{i,j} = cov(Dim_i, Dim_j))$ which generates a square matrix. Then, the covariance

matrices, also called transformation matrices, in this specific example become

$$C = \begin{pmatrix} cov\,(AR, AR) & cov(AR, CAR) & cov(AR, CoupTF\,I) \text{ ------------ } cov(AR, VDR) \\ cov\,(CAR, AR) & cov(CAR, CAR) \text{ --------------------------------- } cov(CAR, VDR) \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ cov(VDR, AR) & cov(VDR, CAR) \text{ --------------------------------- } cov(VDR, VDR) \end{pmatrix}$$

Note that the covariance of the standardized values equals the Pearson correlations, in that it

provides any increased proportion or decreased relationship between two variables. From

matrix algebraic calculations, there are 50 pairs of eigenvectors and eigenvalues derived.

Briefly, the eigenvectors are, literally, a transforming vector directing position of the

transformation matrix on the coordinate plane when multiplied on the left of the vector.

Several features of eigenvectors should be mentioned: they 1) are only found for square matrices, 2) have the same number of eigenvectors as number of dimensions (for example, 50 eigenvectors in 50 dimensions because it is 50 x 50 matrix), 3) are all perpendicular to each other regardless of how many dimensions there are. Thus, fifty pairs of eigenvalues and eigenvectors appeared in this specific study. When eigenvectors are changed to unit vectors, which is a necessary step in PCA, it provides 50 different eigenvalues on each 50 unit vector as follows;

$$e_1 X \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ \vdots \\ \vdots \\ A_{49} \\ A_{50} \end{pmatrix} \text{-------------} e_{27} X \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ \vdots \\ \vdots \\ A_{49} \\ A_{50} \end{pmatrix} \text{-------------} e_{50} X \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ \vdots \\ \vdots \\ A_{49} \\ A_{50} \end{pmatrix}$$

where e represents eigenvalue and $A_1$ through $A_{50}$ represent individual components of each unit eigenvector. As mentioned above, it is important to note that the eigenvector with the highest eigenvalue represents the principal component dataset which provides the most significant relationship between the data and dimensions. In other words, the data point or NRs (in this research) in the first PCA is ideally able to make best explanation for the distribution of all covariances. In general, this is not case so that the combined group of the first several PCAs is generally utilized to explain the covariance distributions. Next, a feature vector is generated combining the individual PCAs in the order of significance from highest to lowest eigenvalues. Thus, it can be written as *Feature Vector* = (eig1 eig2 eig3 … eig50) where eig represents an eigenvector. In this thesis, the first 4 eigenvalues (e.g. first 4 PCAs)

were chosen because the cumulative value is 0.62, meaning 62% of the explanation of all features in this lung study are described by these four eigenvalues. Final data will be generated by

Final Data = Row feature vector x row data adjust,

where the row feature vector is the matrix having the eigenvector components transposed in row from column, with the first PCA positioned on top row and second one next row, etc. The row data adjust is also transposed, mean-adjusted data that each column stands for each data point, thus each row represents a different dimension.


8.3  Bioinformatics

The complete sequence of the whole human genome, along with genome projects of other species, provided important insights of the complexity in genetic regulation. In addition, this genetic information further accelerated biomedical and biological research along with the development of the microarray or DNA chip initiated in the early 90's. In fact, the DNA microarray revolutionized molecular biology, creating a new paradigm of 'functional genomics', which has been applied to all biological fields, structural genomics, proteomics, developmental genomics, and so forth. Functional genomics is attributing to understand the huge amount of genetic information generated from whole genomic sequencing. In comparison to traditional molecular biology in which a research simply plays 'treasure hunting' or performs 'whodunit' analyses, the potential of functional genomics as a tool for molecular biology has been attributed to the development of DNA microarray to monitor for genetic signatures as a whole for cellular physiological changes. Bioinformatics

involves mostly computational biology (*in silico* biology) on vast amounts of experimental data from DNA microarray experiments. Furthermore, the tools are used to efficiently organize new databases, where a flood of information is continuously accumulated, and to build new tools that provide compatibility with the previous database. Although many useful bioinformatics tools are available, here I introduced two tools extensively utilized in this thesis.

8.3.1 Microarray

There are several types of microarray based on the experimental purpose: DNA microarray for detection of transcriptional changes, tissue microarray for immunohistochemistry with multiple tissues, comparative genomic hybridization (CGH) array for detection of chromosomal amplification or deletion. To detect changes in mRNA expression, the most widely used types of arrays include cDNA arrays and oligoarrays. The basic concepts of both types of arrays are considerably different in several aspects, although both pursue detection of differential mRNA expression between samples. First, cDNA arrays have various lengths (several tens of nucleotides up to 5000 nucleotides) of probes whereas oligoarrays include probes with a fixed nucleotide sequence length, i.e., 25 nucleotides for the Affymetrix and 50 nucleotides for the Ilumina array which have been the most widely utilized. There are also sample labeling methods differently adopted to these two types of arrays. Since the cDNA array employs a method to label two samples with different dyes Cy3 and Cy5, it can efficiently compare two different cell lines, for example, without or with knock-down of a specific nuclear receptor to understand its transcriptional targets. Although

it is useful for direct comparison, different hybridization conditions may cause highly variable results, leading to the significantly reduced reproducibility of the data as well as incomparability of two independent experiments. However, the oligoarrays such as the Affymetrix chip and the Ilumina array adopt a single labeling method using biotinylation due to the short length of the probes. This efficiently reduces the weakness of false positive signals in the cDNA array caused by the mismatch hybridization. Also, the reproducibility of the data is highly improved because of the well-defined protocol which makes independent experiments comparable.

Regarding the Affymetrix array, there are several versions of the chips as shown in Table 8.1. There are multiple probesets covering different parts of unique genes. This strategy potentially lowers experimental loss of signal but increases signal variations among probesets for the gene of interest. In addition, recent BLAST analysis, extensively performed by Dr. Luc Girard, revealed that only 63% of the probesets on the affychip match with the Blast results without any cross-hybridization. One percent match genes different from Affymetrix-identified gene and 6% show BLAST genes but no Affymetrix-identified genes. Most importantly, 14% of the probes had no BLAST result even though they were identified by Affymetrix. To our surprise, 16% were unidentified by both Affymetrix and BLAST. Lastly a total of 3% of the probesets including 1% of splice variants can cross-hybridize to other genes, potentially leading to false positive signals. Note that the labeling protocols for Affymetrix arrays are shown in Figure 8.1. A recent appearance of the Ilumina array improved some weakness of the Affymetrix chip discussed above. First, it employs a single and longer probeset (~ 50 nucleotides) compared to the Affychip and has about 30 copies of

each probe, greatly contributing to lower variations in signals and less background. Second, it

reduces the experimental cost to one half of the Affymetrix chip per sample.

In spite of its high-throughput capacity for monitoring genetic signatures on a whole

genome scale, the microarray is qualitative rather than quantitative. Thus one should employ

a secondary experiment for quantitative analysis such as a qRT-PCR assay. Moreover, the

reproducibility also remains to be of concern due to high technical sensitivity of this

technology. Through this whole thesis, both types of arrays were employed to generate data

of genetic signatures in lung cell and tissue panels, and were analyzed by biostatistics and

bioinformatics analysis using the following software.

8.3.2. Matrix and Eisen

The Matrix software, developed by Dr. Luc Girard, is a multipurpose program

including various types of templates for analyzing several types of DNA microarrays such as

CGH, Affymetrix, Ilumina array, and Oligoarray. Each template performs text file formatting

of individual experiments into manageable datasets with signals (or detection calls) with p-

value, and annotation information including gene name, gene ontology for molecular and

biological annotations, LocusLink, chromosomal position, etc. In addition, it also includes

information on the degree of matching of the 11 probe sets for the individual genes from the

BLAST analysis discussed above. Once the text file is applied to the compatible template

selected from a pull-down menu, the dataset is normalized and includes all information

mentioned above. The valid Affymetrix files with the normalization process is denoted with

the extension '.mtx', a read only Excel file. Then, a package of bioinformatics tools can be

utilized according to the various analytic purposes;

1) Color display

This analysis converts the numerical signals (values in $\log_2$) into the color code in

a blue scale for visualization where the darker blues represent the higher

expression signals, whereas the gray scale corresponds to all absent signals. There

is a default setting of 0.065 which is the cutoff threshold p-value for defining

Present call from Absent call.

2) Log ratio for individual genes between two samples

This analysis is the most useful method to find any set of genes, in any preference

fold difference (2 fold or more than), differentially expressed between two

corresponding samples.

3) Combine replicate genes

Due to multiple replicate genes on an array (especially Affymetrix array), one

frequently wishes to cluster the same genes together with an average value of

their signals. This tool provides an efficient method for that purpose using the

UniGene IDs by the gene annotations. The results include a number of replicates

in 'No. Accessions' column, and the correlation value between the replicates

appeared in 'Replicate Correlation' column. The Pearson correlations are

calculated between two replicates or each replicate and the combined average.

4) Scatter plot

Based on the calculation of $\log_2$ ratio between the two samples to be compared,

this analysis plots a point for each gene on the coordinate plane where X and Y

value indicate the expression value n the two samples compared. Based upon the

described fold difference (e.g., 2-fold or 4-fold), the scattered dots are colored in

red for higher expression genes and green for lower expression genes. In addition,

the corresponding 'Data' worksheet is provided with the gene list containing

detailed information including annotations.

5) Create a subarray with a partial gene list

This tool is utilized to extract a subset of genes of particular interest and creat a

new '.mtx' file containing only those genes for further analysis. In this thesis, this

function was used for creation of NR subarrays from the main Affymetrix dataset

as in figure 3.2.

6) Correlations with sample properties

This feature of bioinformatics is extremely useful to get Pearson correlations

between any quantifiable sample characteristics and each gene's expression

profile. For example, qRT-PCR data of an individual NR is correlated to

Affymetrix genes in a panel of 48 lung cells as in Figure 2.2. Thus, the Pearson

correlation matrix is generated between NRs and approximately 30,000 genes.

Following the sort process, genes with statistical significance are listed with a

color code, where red and green indicate positive and negative correlation,

respectively, whereas the dark grey has no correlation. The clustering analysis can

be accompanied along with the correlations.

7) Hierarchical clustering analysis based on Pearson correlation (both centered and
uncentered).

 Following the color display, clustering analysis is performed based on the
calculation of the Pearson correlation (centered or uncentered) with the formula
discussed in section 8.2.2. This function has been most widely used to find a subset
of components with any statistical patterning. I have shown multiple clusters of
NRs potentially relevant to the specific physiologic regulation (Figure 1.8) as well
as the tumorigenic process (Figure 3.9). This provides the best cluster which is
believed to be useful to predict functional relevance between NRs within the same
cluster or provide an insight to the functional role of the specific NR with unknown
function.

Although it includes distinguished features for bioinformatics analysis on various types of
microarray data, the Matrix software does not support analysis of datasets that are expressed
as log ratios of one array compared to another. Thus, next, I briefly introduce the Eisen
software 'Cluster' and 'TreeView' available on **http://rana.lbl.gov/**. As a supplement for the
Matrix program, the Eisen program includes various types of clustering techniques including
hierarchical, k-Means, Self-organizing Map (SOM) and PCA. In this thesis, the Pearson
correlation matrix is generated between any pair of NRs in several panels of samples
including the lung cell panel, tissue panels, and mouse tissue samples. This matrix consists of
both negative and positive correlations, which is followed by hierarchical clustering analysis
using 'Cluster'. Further application of 'TreeView' generates clustering view in various image
formats -- PDF, JPEG, and Bmp -- with color codes, red for positive correlation and green

for negative correlation. Application of both programs is expected to be greatly helpful for

bioinformatics analysis of the vast amount of data.

Table 8.1 Affymetrix CHIP

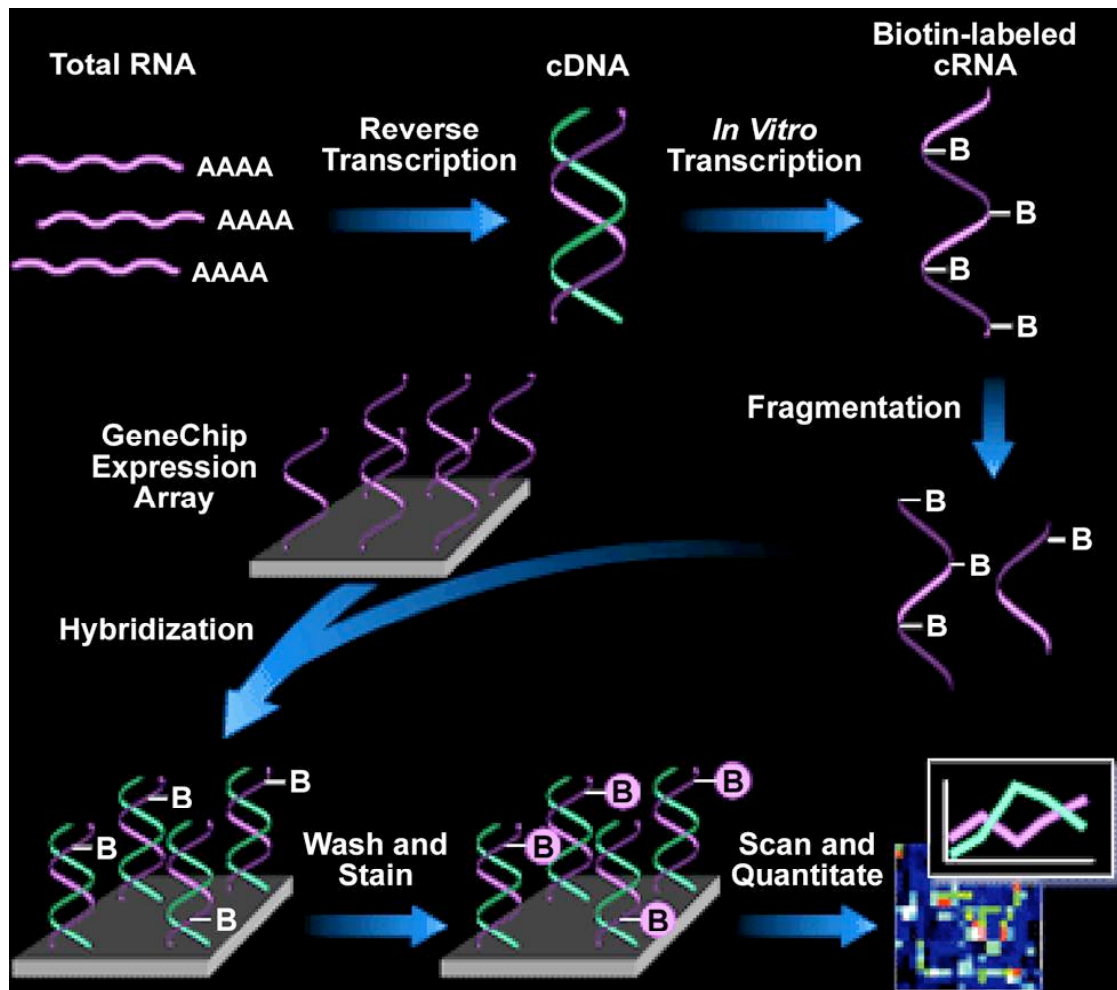| Type of CHIP | No. probesets (No. unique genes) | Detection |
| --- | --- | --- |
| Affymetrix HG-U133A | 22,283 probe sets (12,659 unique genes) | 78% mRNA, 22% EST |
| Affymetrix HG-U133B | 22,645 probe sets (9,147 unique genes) | 20% mRNA, 80% EST |
| Affymetrix HG-U133A and HG-U133B | 44,928 probe sets (18,281 unique genes) | 48% mRNA, 52% EST |
| Affymetrix HG-U133-Plus2 | 54,675 probe sets (21,627 unique genes) | 54% mRNA, 46% EST |

Figure 8.1 Labeling Protocol Affymetrix Chip.

# BIBLIOGRAPHY

Abu, J., Batuwangala, M., Herbert, K., and Symonds, P. (2005). Retinoic acid and retinoid receptors: potential chemopreventive and therapeutic role in cervical cancer. Lancet Oncol *6*, 712-720.

ACS (2006). Cancer Statistics (American Cancer Society).

Aggarwal, B. B., Shishodia, S., Takada, Y., Banerjee, S., Newman, R. A., Bueso-Ramos, C. E., and Price, J. E. (2005). Curcumin suppresses the paclitaxel-induced nuclear factor-kappaB pathway in breast cancer cells and inhibits lung metastasis of human breast cancer in nude mice. Clin Cancer Res *11*, 7490-7498.

Akashi, M., and Takumi, T. (2005). The orphan nuclear receptor RORalpha regulates circadian transcription of the mammalian core-clock Bmal1. Nat Struct Mol Biol *12*, 441-448.

Allred, C. D., and Kilgore, M. W. (2005). Selective activation of PPARgamma in breast, colon, and lung cancer cell lines. Mol Cell Endocrinol *235*, 21-29.

Aung, C. S., Faddy, H. M., Lister, E. J., Monteith, G. R., and Roberts-Thomson, S. J. (2006). Isoform specific changes in PPAR alpha and beta in colon and breast cancer with differentiation. Biochem Biophys Res Commun *340*, 656-660.

Ballard, P. L., and Ballard, R. A. (1974). Cytoplasmic receptor for glucocorticoids in lung of the human fetus and neonate. J Clin Invest *53*, 477-486.

Ballard, P. L., Baxter, J. D., Higgins, S. J., Rousseau, G. G., and Tomkins, G. M. (1974). General presence of glucocorticoid receptors in mammalian tissues. Endocrinology *94*, 998-1002.

Barak, Y., Nelson, M. C., Ong, E. S., Jones, Y. Z., Ruiz-Lozano, P., Chien, K. R., Koder, A., and Evans, R. M. (1999). PPAR gamma is required for placental, cardiac, and adipose tissue development. Mol Cell *4*, 585-595.

Barish, G. D. (2006). Peroxisome proliferator-activated receptors and liver X receptors in atherosclerosis and immunity. J Nutr *136*, 690-694.

Belinsky, S. A. (1998). Role of the cytosine DNA-methyltransferase and p16INK4a genes in the development of mouse lung tumors. Exp Lung Res *24*, 463-479.

Benner, S. E., Lippman, S. M., and Hong, W. K. (1995). Current status of retinoid chemoprevention of lung cancer. Oncology (Williston Park) *9*, 205-210; discussion 210, 213-204, 216.

Berge, K. E., Tian, H., Graf, G. A., Yu, L., Grishin, N. V., Schultz, J., Kwiterovich, P., Shan, B., Barnes, R., and Hobbs, H. H. (2000). Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. Science *290*, 1771-1775.

Bild, A. H., Yao, G., Chang, J. T., Wang, Q., Potti, A., Chasse, D., Joshi, M. B., Harpole, D., Lancaster, J. M., Berchuck, A*., et al.* (2006). Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature *439*, 353-357.

Blaschke, F., Caglayan, E., and Hsueh, W. A. (2006a). Peroxisome proliferator-activated receptor gamma agonists: their role as vasoprotective agents in diabetes. Endocrinol Metab Clin North Am *35*, 561-574, ix.

Blaschke, F., Takata, Y., Caglayan, E., Law, R. E., and Hsueh, W. A. (2006b). Obesity, peroxisome proliferator-activated receptor, and atherosclerosis in type 2 diabetes. Arterioscler Thromb Vasc Biol *26*, 28-40.

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

Botrugno, O. A., Fayard, E., Annicotte, J. S., Haby, C., Brennan, T., Wendling, O., Tanaka, T., Kodama, T., Thomas, W., Auwerx, J., and Schoonjans, K. (2004). Synergy between LRH-1 and beta-catenin induces G1 cyclin-mediated cell proliferation. Mol Cell *15*, 499-509.

Bren-Mattison, Y., Van Putten, V., Chan, D., Winn, R., Geraci, M. W., and Nemenoff, R. A. (2005). Peroxisome proliferator-activated receptor-gamma (PPAR(gamma)) inhibits tumorigenesis by reversing the undifferentiated phenotype of metastatic non-small-cell lung cancer cells (NSCLC). Oncogene *24*, 1412-1422.

Casey, G. (1997). The BRCA1 and BRCA2 breast cancer genes. Curr Opin Oncol *9*, 88-93.

Cesario, R. M., Stone, J., Yen, W. C., Bissonnette, R. P., and Lamph, W. W. (2006). Differentiation and growth inhibition mediated via the RXR:PPARgamma heterodimer in colon cancer. Cancer Lett *240*, 225-233.

Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001). Nuclear receptors and lipid physiology: opening the X-files. Science *294*, 1866-1870.

Chen, A. C., and Petrylak, D. P. (2004). Complications of androgen deprivation therapy in men with prostate cancer. Curr Oncol Rep *6*, 209-215.

Chen, H. Y., Yu, S. L., Chen, C. H., Chang, G. C., Chen, C. Y., Yuan, A., Cheng, C. L., Wang, C. H., Terng, H. J., Kao, S. F.*, et al.* (2007). A five-gene signature and clinical outcome in non-small-cell lung cancer. N Engl J Med *356*, 11-20.

Cheng, L., Ding, G., Qin, Q., Huang, Y., Lewis, W., He, N., Evans, R. M., Schneider, M. D., Brako, F. A., Xiao, Y.*, et al.* (2004). Cardiomyocyte-restricted peroxisome proliferator-activated receptor-delta deletion perturbs myocardial fatty acid oxidation and leads to cardiomyopathy. Nat Med *10*, 1245-1250.

Chmelar, R., Buchanan, G., Need, E. F., Tilley, W., and Greenberg, N. M. (2007). Androgen receptor coregulators and their involvement in the development and progression of prostate cancer. Int J Cancer *120*, 719-733.

Choi, M., Moschetta, A., Bookout, A. L., Peng, L., Umetani, M., Holmstrom, S. R., Suino-Powell, K., Xu, H. E., Richardson, J. A., Gerard, R. D.*, et al.* (2006). Identification of a hormonal basis for gallbladder filling. Nat Med *12*, 1253-1255.

Choy, L., and Derynck, R. (2003). Transforming growth factor-beta inhibits adipocyte differentiation by Smad3 interacting with CCAAT/enhancer-binding protein (C/EBP) and repressing C/EBP transactivation function. J Biol Chem *278*, 9609-9619.

Costet, P., Luo, Y., Wang, N., and Tall, A. R. (2000). Sterol-dependent transactivation of the ABC1 promoter by the liver X receptor/retinoid X receptor. J Biol Chem *275*, 28240-28245.

Crawford, Y. G., Gauthier, M. L., Joubel, A., Mantei, K., Kozakiewicz, K., Afshari, C. A., and Tlsty, T. D. (2004). Histologically normal human mammary epithelia with silenced p16(INK4a) overexpress COX-2, promoting a premalignant program. Cancer Cell *5*, 263-273.

de Rosny, E., de Groot, A., Jullian-Binard, C., Gaillard, J., Borel, F., Pebay-Peyroula, E., Fontecilla-Camps, J. C., and Jouve, H. M. (2006). Drosophila nuclear receptor E75 is a thiolate hemoprotein. Biochemistry *45*, 9727-9734.

De Servi, B., Hermani, A., Medunjanin, S., and Mayer, D. (2005). Impact of PKCdelta on estrogen receptor localization and activity in breast cancer cells. Oncogene *24*, 4946-4955.

de Vos, D. (2005). Epigenetic drugs: a longstanding story. Semin Oncol *32*, 437-442.

Debes, J. D., and Tindall, D. J. (2004). Mechanisms of androgen-refractory prostate cancer. N Engl J Med *351*, 1488-1490.

Encyclopedia, W., In Wilkipedia Encyclopedia.

Fu, M., Sun, T., Bookout, A. L., Downes, M., Yu, R. T., Evans, R. M., and Mangelsdorf, D. J. (2005). A Nuclear Receptor Atlas: 3T3-L1 adipogenesis. Mol Endocrinol *19*, 2437-2450.

Fukumoto, K., Yano, Y., Virgona, N., Hagiwara, H., Sato, H., Senba, H., Suzuki, K., Asano, R., Yamada, K., and Yano, T. (2005). Peroxisome proliferator-activated receptor delta as a molecular target to regulate lung cancer cell growth. FEBS Lett *579*, 3829-3836.

Garcia-Manero, G., Kantarjian, H. M., Sanchez-Gonzalez, B., Yang, H., Rosner, G., Verstovsek, S., Rytting, M., Wierda, W. G., Ravandi, F., Koller, C.*, et al.* (2006). Phase I/II study of the combination of 5-aza-2' -deoxycytidine with valproic acid in patients with leukemia. Blood.

Gazdar, A. F. (2003). Environmental tobacco smoke, carcinogenesis, and angiogenesis: a double whammy? Cancer Cell *4*, 159-160.

Gazdar, A. F., and Carbone, M. (2003). Molecular pathogenesis of malignant mesothelioma and its relationship to simian virus 40. Clin Lung Cancer *5*, 177-181.

Giguere, V., Yang, N., Segui, P., and Evans, R. M. (1988). Identification of a new class of steroid hormone receptors. Nature *331*, 91-94.

Goodwin, B., Jones, S. A., Price, R. R., Watson, M. A., McKee, D. D., Moore, L. B., Galardi, C., Wilson, J. G., Lewis, M. C., Roth, M. E.*, et al.* (2000). A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. Mol Cell *6*, 517-526.

Goyal, P., Weissmann, N., Grimminger, F., Hegel, C., Bader, L., Rose, F., Fink, L., Ghofrani, H. A., Schermuly, R. T., Schmidt, H. H.*, et al.* (2004). Upregulation of NAD(P)H oxidase 1 in hypoxia activates hypoxia-inducible factor 1 via increase in reactive oxygen species. Free Radic Biol Med *36*, 1279-1288.

Grommes, C., Landreth, G. E., and Heneka, M. T. (2004). Antineoplastic effects of peroxisome proliferator-activated receptor gamma agonists. Lancet Oncol *5*, 419-429.

Gu, P., Goodwin, B., Chung, A. C., Xu, X., Wheeler, D. A., Price, R. R., Galardi, C., Peng, L., Latour, A. M., Koller, B. H.*, et al.* (2005). Orphan nuclear receptor LRH-1 is

required to maintain Oct4 expression at the epiblast stage of embryonic development. Mol Cell Biol *25*, 3492-3505.

Gupta, R. A., Sarraf, P., Mueller, E., Brockman, J. A., Prusakiewicz, J. J., Eng, C., Willson, T. M., and DuBois, R. N. (2003). Peroxisome proliferator-activated receptor gamma-mediated differentiation: a mutation in colon cancer cells reveals divergent and cell type-specific mechanisms. J Biol Chem *278*, 22669-22677.

Han, S., Ritzenthaler, J. D., Wingerd, B., and Roman, J. (2005). Activation of peroxisome proliferator-activated receptor beta/delta (PPARbeta/delta) increases the expression of prostaglandin E2 receptor subtype EP4. The roles of phosphatidylinositol 3-kinase and CCAAT/enhancer-binding protein beta. J Biol Chem *280*, 33240-33249.

He, W., Barak, Y., Hevener, A., Olson, P., Liao, D., Le, J., Nelson, M., Ong, E., Olefsky, J. M., and Evans, R. M. (2003). Adipose-specific peroxisome proliferator-activated receptor gamma knockout causes insulin resistance in fat and liver but not in muscle. Proc Natl Acad Sci U S A *100*, 15712-15717.

Hershberger, P. A., Vasquez, A. C., Kanterewicz, B., Land, S., Siegfried, J. M., and Nichols, M. (2005). Regulation of endogenous gene expression in human non-small cell lung cancer cells by estrogen receptor ligands. Cancer Res *65*, 1598-1605.

Hevener, A. L., He, W., Barak, Y., Le, J., Bandyopadhyay, G., Olson, P., Wilkes, J., Evans, R. M., and Olefsky, J. (2003). Muscle-specific Pparg deletion causes insulin resistance. Nat Med *9*, 1491-1497.

Heyman, R. A., Mangelsdorf, D. J., Dyck, J. A., Stein, R. B., Eichele, G., Evans, R. M., and Thaller, C. (1992). 9-cis retinoic acid is a high affinity ligand for the retinoid X receptor. Cell *68*, 397-406.

Horkko, T. T., Tuppurainen, K., George, S. M., Jernvall, P., Karttunen, T. J., and Makinen, M. J. (2006). Thyroid hormone receptor beta1 in normal colon and colorectal cancer-association with differentiation, polypoid growth type and K-ras mutations. Int J Cancer *118*, 1653-1659.

Hortobagyi, G. N., Hung, M. C., and Lopez-Berestein, G. (1998). A Phase I multicenter study of E1A gene therapy for patients with metastatic breast cancer and epithelial ovarian cancer that overexpresses HER-2/neu or epithelial ovarian cancer. Hum Gene Ther *9*, 1775-1798.

http://www.cancer.org The History of Cancer.

HTTP://WWW.NURSA.ORG. Nuclear Receptor.

http://www.rare-cancer.org History of Cancer.

http://www.statsoft.com/textbook/stbasic.html.

Hussain, E. A., Mehta, R. R., Ray, R., Das Gupta, T. K., and Mehta, R. G. (2003). Efficacy and mechanism of action of 1alpha-hydroxy-24-ethyl-cholecalciferol (1alpha[OH]D5) in breast cancer prevention and therapy. Recent Results Cancer Res *164*, 393-411.

Inagaki, T., Choi, M., Moschetta, A., Peng, L., Cummins, C. L., McDonald, J. G., Luo, G., Jones, S. A., Goodwin, B., Richardson, J. A*., et al.* (2005). Fibroblast growth factor 15 functions as an enterohepatic signal to regulate bile acid homeostasis. Cell Metab *2*, 217-225.

Jalving, M., Koornstra, J. J., De Jong, S., De Vries, E. G., and Kleibeuker, J. H. (2005). Review article: the potential of combinational regimen with non-steroidal anti-inflammatory drugs in the chemoprevention of colorectal cancer. Aliment Pharmacol Ther *21*, 321-339.

Janowski, B. A., Willy, P. J., Devi, T. R., Falck, J. R., and Mangelsdorf, D. J. (1996). An oxysterol signalling pathway mediated by the nuclear receptor LXR alpha. Nature *383*, 728-731.

Jansson, E. A., Are, A., Greicius, G., Kuo, I. C., Kelly, D., Arulampalam, V., and Pettersson, S. (2005). The Wnt/beta-catenin signaling pathway targets PPARgamma activity in colon cancer cells. Proc Natl Acad Sci U S A *102*, 1460-1465.

Jarmalaite, S., Kannio, A., Anttila, S., Lazutka, J. R., and Husgafvel-Pursiainen, K. (2003). Aberrant p16 promoter methylation in smokers and former smokers with nonsmall cell lung cancer. Int J Cancer *106*, 913-918.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R. T., Crowley, D., Tuveson, D. A., and Jacks, T. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. Nature *410*, 1111-1116.

Joly-Pharaboz, M. O., Ruffion, A., Roch, A., Michel-Calemard, L., Andre, J., Chantepie, J., Nicolas, B., and Panaye, G. (2000). Inhibition of growth and induction of apoptosis by androgens of a variant of LNCaP cell line. J Steroid Biochem Mol Biol *73*, 237-249.

Jordan, V. C. (2002). The secrets of selective estrogen receptor modulation: cell-specific coregulation. Cancer Cell *1*, 215-217.

Jordan, V. C. (2004). Selective estrogen receptor modulation: concept and consequences in cancer. Cancer Cell *5*, 207-213.

Joseph, S. B., Bradley, M. N., Castrillo, A., Bruhn, K. W., Mak, P. A., Pei, L., Hogenesch, J., O'Connell R, M., Cheng, G., Saez, E*., et al.* (2004). LXR-dependent gene expression is important for macrophage survival and the innate immune response. Cell *119*, 299-309.

Kersten, S., Seydoux, J., Peters, J. M., Gonzalez, F. J., Desvergne, B., and Wahli, W. (1999). Peroxisome proliferator-activated receptor alpha mediates the adaptive response to fasting. J Clin Invest *103*, 1489-1498.

Keshamouni, V. G., Reddy, R. C., Arenberg, D. A., Joel, B., Thannickal, V. J., Kalemkerian, G. P., and Standiford, T. J. (2004). Peroxisome proliferator-activated receptor-gamma activation inhibits tumor progression in non-small-cell lung cancer. Oncogene *23*, 100-108.

Khuri, F. R., and Lippman, S. M. (2000). Lung cancer chemoprevention. Semin Surg Oncol *18*, 100-105.

Kim, B. N., Yamamoto, H., Ikeda, K., Damdinsuren, B., Sugita, Y., Ngan, C. Y., Fujie, Y., Ogawa, M., Hata, T., Ikeda, M*., et al.* (2005a). Methylation and expression of p16INK4 tumor suppressor gene in primary colorectal cancer tissues. Int J Oncol *26*, 1217-1226.

Kim, C. F., Jackson, E. L., Woolfenden, A. E., Lawrence, S., Babar, I., Vogel, S., Crowley, D., Bronson, R. T., and Jacks, T. (2005b). Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell *121*, 823-835.

Kim, D. H., Kim, J. S., Ji, Y. I., Shim, Y. M., Kim, H., Han, J., and Park, J. (2003). Hypermethylation of RASSF1A promoter is associated with the age at starting smoking and a poor prognosis in primary non-small cell lung cancer. Cancer Res *63*, 3743-3746.

Kim, D. H., Nelson, H. H., Wiencke, J. K., Zheng, S., Christiani, D. C., Wain, J. C., Mark, E. J., and Kelsey, K. T. (2001). p16(INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. Cancer Res *61*, 3419-3424.

Kim, J. S., Kim, J. W., Han, J., Shim, Y. M., Park, J., and Kim, D. H. (2006). Cohypermethylation of p16 and FHIT promoters as a prognostic factor of recurrence in surgically resected stage I non-small cell lung cancer. Cancer Res *66*, 4049-4054.

Kishimoto, M., Kohno, T., Okudela, K., Otsuka, A., Sasaki, H., Tanabe, C., Sakiyama, T., Hirama, C., Kitabayashi, I., Minna, J. D*., et al.* (2005). Mutations and deletions of the CBP gene in human lung cancer. Clin Cancer Res *11*, 512-519.

Knapp, P., Jarzabek, K., Blachnio, A., and Zbroch, T. (2006). [The role of peroxisome proliferator-activated receptors (PPAR) in carcinogenesis]. Ginekol Pol *77*, 643-651.

Kopp, P. (1998). Resveratrol, a phytoestrogen found in red wine. A possible explanation for the conundrum of the 'French paradox'? Eur J Endocrinol *138*, 619-620.

Laffitte, B. A., Repa, J. J., Joseph, S. B., Wilpitz, D. C., Kast, H. R., Mangelsdorf, D. J., and Tontonoz, P. (2001). LXRs control lipid-inducible expression of the apolipoprotein E gene in macrophages and adipocytes. Proc Natl Acad Sci U S A *98*, 507-512.

Lee, C. H., and Plutzky, J. (2006). Liver X receptor activation and high-density lipoprotein biology: a reversal of fortune? Circulation *113*, 5-8.

Lee, C. T., Li, L., Takamoto, N., Martin, J. F., Demayo, F. J., Tsai, M. J., and Tsai, S. Y. (2004a). The nuclear orphan receptor COUP-TFII is required for limb and skeletal muscle development. Mol Cell Biol *24*, 10835-10843.

Lee, H. K., Lee, Y. K., Park, S. H., Kim, Y. S., Park, S. H., Lee, J. W., Kwon, H. B., Soh, J., Moore, D. D., and Choi, H. S. (1998). Structure and expression of the orphan nuclear receptor SHP gene. J Biol Chem *273*, 14398-14402.

Lee, H. Y., Chang, Y. S., Han, J. Y., Liu, D. D., Lee, J. J., Lotan, R., Spitz, M. R., and Hong, W. K. (2005a). Effects of 9-cis-retinoic acid on the insulin-like growth factor axis in former smokers. J Clin Oncol *23*, 4439-4449.

Lee, H. Y., Moon, H., Chun, K. H., Chang, Y. S., Hassan, K., Ji, L., Lotan, R., Khuri, F. R., and Hong, W. K. (2004b). Effects of insulin-like growth factor binding protein-3 and farnesyltransferase inhibitor SCH66336 on Akt expression and apoptosis in non-small-cell lung cancer cells. J Natl Cancer Inst *96*, 1536-1548.

Lee, K. W., Ma, L., Yan, X., Liu, B., Zhang, X. K., and Cohen, P. (2005b). Rapid apoptosis induction by IGFBP-3 involves an insulin-like growth factor-independent nucleomitochondrial translocation of RXRalpha/Nur77. J Biol Chem *280*, 16942-16948.

Lehmann, J. M., Kliewer, S. A., Moore, L. B., Smith-Oliver, T. A., Oliver, B. B., Su, J. L., Sundseth, S. S., Winegar, D. A., Blanchard, D. E., Spencer, T. A., and Willson, T. M. (1997). Activation of the nuclear receptor LXR by oxysterols defines a new hormone response pathway. J Biol Chem *272*, 3137-3140.

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

Link, K. H., Sagban, T. A., Morschel, M., Tischbirek, K., Holtappels, M., Apell, V., Zayed, K., Kornmann, M., and Staib, L. (2005). Colon cancer: survival after curative surgery. Langenbecks Arch Surg *390*, 83-93.

Lippman, S. M., Benner, S. E., and Hong, W. K. (1993). Chemoprevention strategies in lung carcinogenesis. Chest *103*, 15S-19S.

Liu, S., Sugimoto, Y., Kulp, S. K., Jiang, J., Chang, H. L., Park, K. Y., Kashida, Y., and Lin, Y. C. (2002). Estrogenic down-regulation of protein tyrosine phosphatase gamma (PTP gamma) in human breast is associated with estrogen receptor alpha. Anticancer Res *22*, 3917-3923.

Liu, X., Pisha, E., Tonetti, D. A., Yao, D., Li, Y., Yao, J., Burdette, J. E., and Bolton, J. L. (2003). Antiestrogenic and DNA damaging effects induced by tamoxifen and toremifene metabolites. Chem Res Toxicol *16*, 832-837.

Lockless, S. W., and Ranganathan, R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. Science *286*, 295-299.

Lu, H. C., Eichele, G., and Thaller, C. (1997). Ligand-bound RXR can mediate retinoid signal transduction during embryogenesis. Development *124*, 195-203.

Lu, T. T., Makishima, M., Repa, J. J., Schoonjans, K., Kerr, T. A., Auwerx, J., and Mangelsdorf, D. J. (2000). Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. Mol Cell *6*, 507-515.

Luo, Y., and Tall, A. R. (2000). Sterol upregulation of human CETP expression in vitro and in transgenic mice by an LXR element. J Clin Invest *105*, 513-520.

Lynch, T., Jr., and Kim, E. (2005). Optimizing chemotherapy and targeted agent combinations in NSCLC. Lung Cancer *50S2*, S25-S32.

Majumdar, A. P., Kodali, U., and Jaszewski, R. (2004). Chemopreventive role of folic acid in colorectal cancer. Front Biosci *9*, 2725-2732.

Makishima, M., Lu, T. T., Xie, W., Whitfield, G. K., Domoto, H., Evans, R. M., Haussler, M. R., and Mangelsdorf, D. J. (2002). Vitamin D receptor as an intestinal bile acid sensor. Science *296*, 1313-1316.

Makishima, M., Okamoto, A. Y., Repa, J. J., Tu, H., Learned, R. M., Luk, A., Hull, M. V., Lustig, K. D., Mangelsdorf, D. J., and Shan, B. (1999). Identification of a nuclear receptor for bile acids. Science *284*, 1362-1365.

Mangelsdorf, D. J., Borgmeyer, U., Heyman, R. A., Zhou, J. Y., Ong, E. S., Oro, A. E., Kakizuka, A., and Evans, R. M. (1992). Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. Genes Dev *6*, 329-344.

Marcus, J. N., Watson, P., Page, D. L., Narod, S. A., Tonin, P., Lenoir, G. M., Serova, O., and Lynch, H. T. (1997). BRCA2 hereditary breast cancer pathophenotype. Breast Cancer Res Treat *44*, 275-277.

Marmorstein, R. (2004). Structure and chemistry of the Sir2 family of NAD+-dependent histone/protein deactylases. Biochem Soc Trans *32*, 904-909.

Maruyama, R., Sugio, K., Yoshino, I., Maehara, Y., and Gazdar, A. F. (2004). Hypermethylation of FHIT as a prognostic marker in nonsmall cell lung carcinoma. Cancer *100*, 1472-1477.

Mehta, R. G., and Mehta, R. R. (2002). Vitamin D and cancer. J Nutr Biochem *13*, 252-264.

Meuwissen, R., Linn, S. C., Linnoila, R. I., Zevenhoven, J., Mooi, W. J., and Berns, A. (2003). Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell *4*, 181-189.

Minna, J. D. (1993). The molecular biology of lung cancer pathogenesis. Chest *103*, 449S-456S.

Minna, J. D. (2005). HARRISON'S Principles of INTERNAL MEDICINE, 15th Edition edn: McGraw-Hill).

Minna, J. D., Kurie, J. M., and Jacks, T. (2003). A big step in the study of small cell lung cancer. Cancer Cell *4*, 163-166.

Mitchell, C. (1992). Multistage carcinogenesis in paediatric and adult cancers. Eur J Cancer *28*, 296-298.

Moreno, P. R., Purushothaman, K. R., Sirol, M., Levy, A. P., and Fuster, V. (2006). Neovascularization in human atherosclerosis. Circulation *113*, 2245-2252.

Morriss-Kay, G. M., and Ward, S. J. (1999). Retinoids and mammalian development. Int Rev Cytol *188*, 73-131.

Motola, D. L., Cummins, C. L., Rottiers, V., Sharma, K. K., Li, T., Li, Y., Suino-Powell, K., Xu, H. E., Auchus, R. J., Antebi, A., and Mangelsdorf, D. J. (2006). Identification of ligands for DAF-12 that govern dauer formation and reproduction in C. elegans. Cell *124*, 1209-1223.

Nabha, S. M., Glaros, S., Hong, M., Lykkesfeldt, A. E., Schiff, R., Osborne, K., and Reddy, K. B. (2005). Upregulation of PKC-delta contributes to antiestrogen resistance in mammary tumor cells. Oncogene *24*, 3166-3176.

Nakagawa, K., Kawaura, A., Kato, S., Takeda, E., and Okano, T. (2005a). 1 alpha,25-Dihydroxyvitamin D(3) is a preventive factor in the metastasis of lung cancer. Carcinogenesis *26*, 429-440.

Nakagawa, K., Sasaki, Y., Kato, S., Kubodera, N., and Okano, T. (2005b). 22-Oxa-1alpha,25-dihydroxyvitamin D3 inhibits metastasis and angiogenesis in lung cancer. Carcinogenesis *26*, 1044-1054.

Norris, A. W., Chen, L., Fisher, S. J., Szanto, I., Ristow, M., Jozsi, A. C., Hirshman, M. F., Rosen, E. D., Goodyear, L. J., Gonzalez, F. J.*, et al.* (2003). Muscle-specific PPARgamma-deficient mice develop increased adiposity and insulin resistance but respond to thiazolidinediones. J Clin Invest *112*, 608-618.

Okamoto, M., Lee, C., and Oyasu, R. (1997). Autocrine effect of androgen on proliferation of an androgen responsive prostatic carcinoma cell line, LNCAP: role of interleukin-6. Endocrinology *138*, 5071-5074.

Osipo, C., Gajdos, C., Liu, H., Chen, B., and Jordan, V. C. (2003). Paradoxical action of fulvestrant in estradiol-induced regression of tamoxifen-stimulated breast cancer. J Natl Cancer Inst *95*, 1597-1608.

Park, S. Y., Kim, Y. J., Gao, A. C., Mohler, J. L., Onate, S. A., Hidalgo, A. A., Ip, C., Park, E. M., Yoon, S. Y., and Park, Y. M. (2006). Hypoxia increases androgen receptor activity in prostate cancer cells. Cancer Res *66*, 5121-5129.

Parks, D. J., Blanchard, S. G., Bledsoe, R. K., Chandra, G., Consler, T. G., Kliewer, S. A., Stimmel, J. B., Willson, T. M., Zavacki, A. M., Moore, D. D., and Lehmann, J. M. (1999). Bile acids: natural ligands for an orphan nuclear receptor. Science *284*, 1365-1368.

Pascual, G., Fong, A. L., Ogawa, S., Gamliel, A., Li, A. C., Perissi, V., Rose, D. W., Willson, T. M., Rosenfeld, M. G., and Glass, C. K. (2005). A SUMOylation-dependent pathway mediates transrepression of inflammatory response genes by PPAR-gamma. Nature *437*, 759-763.

Peet, D. J., Janowski, B. A., and Mangelsdorf, D. J. (1998a). The LXRs: a new class of oxysterol receptors. Curr Opin Genet Dev *8*, 571-575.

Peet, D. J., Turley, S. D., Ma, W., Janowski, B. A., Lobaccaro, J. M., Hammer, R. E., and Mangelsdorf, D. J. (1998b). Cholesterol and bile acid metabolism are impaired in mice lacking the nuclear oxysterol receptor LXR alpha. Cell *93*, 693-704.

Petrylak, D. P., Tangen, C. M., Hussain, M. H., Lara, P. N., Jr., Jones, J. A., Taplin, M. E., Burch, P. A., Berry, D., Moinpour, C., Kohli, M.*, et al.* (2004). Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer. N Engl J Med *351*, 1513-1520.

Petty, W. J., Li, N., Biddle, A., Bounds, R., Nitkin, C., Ma, Y., Dragnev, K. H., Freemantle, S. J., and Dmitrovsky, E. (2005). A novel retinoic acid receptor beta isoform and retinoid resistance in lung carcinogenesis. J Natl Cancer Inst *97*, 1645-1651.

Piccart-Gebhart, M. J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., Gianni, L., Baselga, J., Bell, R., Jackisch, C.*, et al.* (2005). Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med *353*, 1659-1672.

Potti, A., Mukherjee, S., Petersen, R., Dressman, H. K., Bild, A., Koontz, J., Kratzke, R., Watson, M. A., Kelley, M., Ginsburg, G. S.*, et al.* (2006). A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. N Engl J Med *355*, 570-580.

Prueitt, R. L., Boersma, B. J., Howe, T. M., Goodman, J. E., Thomas, D. D., Ying, L., Pfiester, C. M., Yfantis, H. G., Cottrell, J. R., Lee, D. H.*, et al.* (2007). Inflammation and IGF-I activate the Akt pathway in breast cancer. Int J Cancer *120*, 796-805.

Ramachandran, U., Kumar, R., and Mittal, A. (2006). Fine tuning of PPAR ligands for type 2 diabetes and metabolic syndrome. Mini Rev Med Chem *6*, 563-573.

Ramirez, R. D., Sheridan, S., Girard, L., Sato, M., Kim, Y., Pollack, J., Peyton, M., Zou, Y., Kurie, J. M., Dimaio, J. M.*, et al.* (2004). Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. Cancer Res *64*, 9027-9034.

Repa, J. J., Berge, K. E., Pomajzl, C., Richardson, J. A., Hobbs, H., and Mangelsdorf, D. J. (2002). Regulation of ATP-binding cassette sterol transporters ABCG5 and ABCG8 by the liver X receptors alpha and beta. J Biol Chem *277*, 18793-18800.

Repa, J. J., Turley, S. D., Lobaccaro, J. A., Medina, J., Li, L., Lustig, K., Shan, B., Heyman, R. A., Dietschy, J. M., and Mangelsdorf, D. J. (2000). Regulation of absorption and ABC1-mediated efflux of cholesterol by RXR heterodimers. Science *289*, 1524-1529.

Rizzo, G., and Fiorucci, S. (2006). PPARs and other nuclear receptors in inflammation. Curr Opin Pharmacol *6*, 421-427.

Rodrigues, S., Bruyneel, E., Rodrigue, C. M., Shahin, E., and Gespach, C. (2004). [Cyclooxygenase 2 and carcinogenesis]. Bull Cancer *91 Spec No,* S61-76.

Rosen, E. D., and Spiegelman, B. M. (2000). Molecular regulation of adipogenesis. Annu Rev Cell Dev Biol *16*, 145-171.

Saez, E., Olson, P., and Evans, R. M. (2003). Genetic deficiency in Pparg does not alter development of experimental prostate cancer. Nat Med *9*, 1265-1266.

Saez, E., Rosenfeld, J., Livolsi, A., Olson, P., Lombardo, E., Nelson, M., Banayo, E., Cardiff, R. D., Izpisua-Belmonte, J. C., and Evans, R. M. (2004). PPAR gamma signaling exacerbates mammary gland tumor development. Genes Dev *18*, 528-540.

Santos-Rosa, H., and Caldas, C. (2005). Chromatin modifier enzymes, the histone code and cancer. Eur J Cancer *41*, 2381-2402.

Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H., and Vennstrom, B. (1986). The c-erb-A protein is a high-affinity receptor for thyroid hormone. Nature *324*, 635-640.

Sarraf, P., Mueller, E., Jones, D., King, F. J., DeAngelo, D. J., Partridge, J. B., Holden, S. A., Chen, L. B., Singer, S., Fletcher, C., and Spiegelman, B. M. (1998). Differentiation and reversal of malignant changes in colon cancer through PPARgamma. Nat Med *4*, 1046-1052.

Sasaki, H., Tanahashi, M., Yukiue, H., Moiriyama, S., Kobayashi, Y., Nakashima, Y., Kaji, M., Kiriyama, M., Fukai, I., Yamakawa, Y., and Fujii, Y. (2002). Decreased perioxisome proliferator-activated receptor gamma gene expression was correlated with poor prognosis in patients with lung cancer. Lung Cancer *36*, 71-76.

Sato, M., Vaughan, M. B., Girard, L., Peyton, M., Lee, W., Shames, D. S., Ramirez, R. D., Sunaga, N., Gazdar, A. F., Shay, J. W., and Minna, J. D. (2006). Multiple oncogenic changes (K-RAS(V12), p53 knockdown, mutant EGFRs, p16 bypass, telomerase) are not sufficient to confer a full malignant phenotype on human bronchial epithelial cells. Cancer Res *66*, 2116-2128.

Schroder, O., Yudina, Y., Sabirsh, A., Zahn, N., Haeggstrom, J. Z., and Stein, J. (2006). 15-deoxy-Delta12,14-prostaglandin J2 inhibits the expression of microsomal prostaglandin E synthase type 2 in colon cancer cells. J Lipid Res *47*, 1071-1080.

Schulman, I. G., Li, C., Schwabe, J. W., and Evans, R. M. (1997). The phantom ligand effect: allosteric control of transcription by the retinoid X receptor. Genes Dev *11*, 299-308.

Seol, W., Hanstein, B., Brown, M., and Moore, D. D. (1998). Inhibition of estrogen receptor action by the orphan receptor SHP (short heterodimer partner). Mol Endocrinol *12*, 1551-1557.

Shao, J., Sheng, H., and DuBois, R. N. (2002). Peroxisome proliferator-activated receptors modulate K-Ras-mediated transformation of intestinal epithelial cells. Cancer Res *62*, 3282-3288.

Shi, X., and Gozani, O. (2005). The fellowships of the INGs. J Cell Biochem *96*, 1127-1136.

Shishodia, S., Potdar, P., Gairola, C. G., and Aggarwal, B. B. (2003). Curcumin (diferuloylmethane) down-regulates cigarette smoke-induced NF-kappaB activation through inhibition of IkappaBalpha kinase in human lung epithelial cells: correlation with suppression of COX-2, MMP-9 and cyclin D1. Carcinogenesis *24*, 1269-1279.

Shulman, A. I., Larson, C., Mangelsdorf, D. J., and Ranganathan, R. (2004). Structural determinants of allosteric ligand activation in RXR heterodimers. Cell *116*, 417-429. Singh, B., Berry, J. A., Shoher, A., Ramakrishnan, V., and Lucci, A. (2005). COX-2 overexpression increases motility and invasion of breast cancer cells. Int J Oncol *26*, 1393-1399.

Sinha, R., Anderson, D. E., McDonald, S. S., and Greenwald, P. (2003). Cancer risk and diet in India. J Postgrad Med *49*, 222-228.

Society, A. C. (2006). Cancer Statistics (American Cancer Society).

Song, L., Yan, W., Deng, M., Song, S., Zhang, J., and Zhao, T. (2004). Aberrations in the fragile histidine triad(FHIT) gene may be involved in lung carcinogenesis in patients with chronic pulmonary tuberculosis. Tumour Biol *25*, 270-275.

Stabile, L. P., Lyker, J. S., Gubish, C. T., Zhang, W., Grandis, J. R., and Siegfried, J. M. (2005). Combined targeting of the estrogen receptor and the epidermal growth factor receptor in non-small cell lung cancer shows enhanced antiproliferative effects. Cancer Res *65*, 1459-1470.

Suzuki, M., Sunaga, N., Shames, D. S., Toyooka, S., Gazdar, A. F., and Minna, J. D. (2004). RNA interference-mediated knockdown of DNA methyltransferase 1 leads to promoter demethylation and gene re-expression in human lung and breast cancer cells. Cancer Res *64*, 3137-3143.

Tannock, I. F., de Wit, R., Berry, W. R., Horti, J., Pluzanska, A., Chi, K. N., Oudard, S., Theodore, C., James, N. D., Turesson, I.*, et al.* (2004). Docetaxel plus prednisone or

mitoxantrone plus prednisone for advanced prostate cancer. N Engl J Med *351*, 1502-1512.

Thomas, S. M., and Grandis, J. R. (2004). Pharmacokinetic and pharmacodynamic properties of EGFR inhibitors under clinical investigation. Cancer Treat Rev *30*, 255-268. Tobin, J. F., and Freedman, L. P. (2006). Nuclear receptors as drug targets in metabolic diseases: new approaches to therapy. Trends Endocrinol Metab *17*, 284-290.

Tontonoz, P., Hu, E., Graves, R. A., Budavari, A. I., and Spiegelman, B. M. (1994). mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. Genes Dev *8*, 1224-1234.

Toulouse, A., Morin, J., Dion, P. A., Houle, B., and Bradley, W. E. (2000). RARbeta2 specificity in mediating RA inhibition of growth of lung cancer-derived cells. Lung Cancer *28*, 127-137.

Umekita, Y., Hiipakka, R. A., Kokontis, J. M., and Liao, S. (1996). Human prostate tumor growth in athymic mice: inhibition by androgens and stimulation by finasteride. Proc Natl Acad Sci U S A *93*, 11802-11807.

Vandoros, G. P., Konstantinopoulos, P. A., Sotiropoulou-Bonikou, G., Kominea, A., Papachristou, G. I., Karamouzis, M. V., Gkermpesi, M., Varakis, I., and Papavassiliou, A. G. (2006). PPAR-gamma is expressed and NF-kB pathway is activated and correlates positively with COX-2 expression in stromal myofibroblasts surrounding colon adenocarcinomas. J Cancer Res Clin Oncol *132*, 76-84.

Virmani, A. K., Rathi, A., Zochbauer-Muller, S., Sacchi, N., Fukuyama, Y., Bryant, D., Maitra, A., Heda, S., Fong, K. M., Thunnissen, F.*, et al.* (2000). Promoter methylation and silencing of the retinoic acid receptor-beta gene in lung carcinomas. J Natl Cancer Inst *92*, 1303-1307.

Wang, D., Wang, H., Shi, Q., Katkuri, S., Walhi, W., Desvergne, B., Das, S. K., Dey, S. K., and DuBois, R. N. (2004a). Prostaglandin E(2) promotes colorectal adenoma growth via transactivation of the nuclear peroxisome proliferator-activated receptor delta. Cancer Cell *6*, 285-295.

Wang, J., Yin, L., and Lazar, M. A. (2006a). The orphan nuclear receptor Rev-erb alpha regulates circadian expression of plasminogen activator inhibitor type 1. J Biol Chem *281*, 33842-33848.

Wang, L. H., Yang, X. Y., Zhang, X., An, P., Kim, H. J., Huang, J., Clarke, R., Osborne, C. K., Inman, J. K., Appella, E., and Farrar, W. L. (2006b). Disruption of estrogen receptor DNA-binding domain and related intramolecular communication restores tamoxifen sensitivity in resistant breast cancer. Cancer Cell *10*, 487-499.

Wang, T., Xu, J., Yu, X., Yang, R., and Han, Z. C. (2006c). Peroxisome proliferator-activated receptor gamma in malignant diseases. Crit Rev Oncol Hematol *58*, 1-14.
Wang, Y., Zhang, Z., Yao, R., Jia, D., Wang, D., Lubet, R. A., and You, M. (2006d). Prevention of lung cancer progression by bexarotene in mouse models. Oncogene *25*, 1320-1329.

Wang, Y. X., Zhang, C. L., Yu, R. T., Cho, H. K., Nelson, M. C., Bayuga-Ocampo, C. R., Ham, J., Kang, H., and Evans, R. M. (2004b). Regulation of muscle fiber type and running endurance by PPARdelta. PLoS Biol *2*, e294.

Weinberger, C., Thompson, C. C., Ong, E. S., Lebo, R., Gruol, D. J., and Evans, R. M. (1986). The c-erb-A gene encodes a thyroid hormone receptor. Nature *324*, 641-646.
Whitsett, T., Carpenter, M., and Lamartiniere, C. A. (2006). Resveratrol, but not EGCG, in the diet suppresses DMBA-induced mammary cancer in rats. J Carcinog *5*, 15.

Wietzke, J. A., Ward, E. C., Schneider, J., and Welsh, J. (2005). Regulation of the human vitamin D3 receptor promoter in breast cancer cells is mediated through Sp1 sites. Mol Cell Endocrinol *230*, 59-68.

Willy, P. J., and Mangelsdorf, D. J. (1997). Unique requirements for retinoid-dependent transcriptional activation by the orphan receptor LXR. Genes Dev *11*, 289-298.
Witton, C. J., Hawe, S. J., Cooke, T. G., and Bartlett, J. M. (2004). Cyclooxygenase 2 (COX2) expression is associated with poor outcome in ER-negative, but not ER-positive, breast cancer. Histopathology *45*, 47-54.

Yamamoto, Y., Moore, R., Goldsworthy, T. L., Negishi, M., and Maronpot, R. R. (2004). The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. Cancer Res *64*, 7197-7200.

Yang, F. G., Zhang, Z. W., Xin, D. Q., Shi, C. J., Wu, J. P., Guo, Y. L., and Guan, Y. F. (2005). Peroxisome proliferator-activated receptor gamma ligands induce cell cycle arrest and apoptosis in human renal carcinoma cell lines. Acta Pharmacol Sin *26*, 753-761.

Yang, X., Downes, M., Yu, R. T., Bookout, A. L., He, W., Straume, M., Mangelsdorf, D. J., and Evans, R. M. (2006). Nuclear receptor expression links the circadian clock to metabolism. Cell *126*, 801-810.

Yin, L., and Lazar, M. A. (2005). The orphan nuclear receptor Rev-erbalpha recruits the N-CoR/histone deacetylase 3 corepressor to regulate the circadian Bmal1 gene. Mol Endocrinol *19*, 1452-1459.

Yin, L., Wang, J., Klein, P. S., and Lazar, M. A. (2006). Nuclear receptor Rev-erbalpha is a critical lithium-sensitive component of the circadian clock. Science *311*, 1002-1005.

Yoo, S. J., Seo, E. J., Lee, J. H., Seo, Y. H., Park, P. W., and Ahn, J. Y. (2006). A complex, four-way variant t(15;17) in acute promyelocytic leukemia. Cancer Genet Cytogenet *167*, 168-171.

Zelcer, N., and Tontonoz, P. (2006). Liver X receptors as integrators of metabolic and inflammatory signaling. J Clin Invest *116*, 607-614.

Zhang, Y., Repa, J. J., Gauthier, K., and Mangelsdorf, D. J. (2001). Regulation of lipoprotein lipase by the oxysterol receptors, LXRalpha and LXRbeta. J Biol Chem *276*, 43018-43024.

Zhu, B. Q., Heeschen, C., Sievers, R. E., Karliner, J. S., Parmley, W. W., Glantz, S. A., and Cooke, J. P. (2003). Second hand smoke stimulates tumor angiogenesis and growth. Cancer Cell *4*, 191-196.

Zhu, W., Wu, B., Zheng, J., Fang, W., Wan, J., and You, J. (1996). Reduced tumorigenicity of metastatic human lung cancer cell subline (PGCL3) transfected with hRAR beta gene. Chin Med Sci J *11*, 13-16.

# VITAE

Yangsik Jeong was born in Hadong, Kyung-sang nam do, Republic of Korea, on December 9, 1970, the first son of Jeom Soon Kim and Ki Hwa Jeong. After graduating from Hak-sung High School in Ulsan, South Korea in 1989, he entered the Kyung-Pook National University in Taegoo. He received the degree of Bachelor of Science in Genetic Engineering from the Kyung-Pook National University in February, 1996. In March 1996, he entered the Graduate School of Life Science at the Pohang University of Science & Technology (POSTECH) where he initiated his work as a M.S. candidate in the laboratory of Dr. Young Chul Sung. From 1996 to 1998, he was employed as a research assistant in the Department of Biology at the POSTECH. He was awarded the degree of Master of Science in viral immunology in July 1998. From September 1998 to December 1999, he was employed as a research scientist in Lucky Biotech Institute, Republic of Korea. From January 2000 to July 2001, he was employed as a research assistant in the MD Anderson Cancer Center, United States of America. In August 2001, he entered the Graduate Program of Biomedical Sciences at the University of Texas Southwestern Medical Center at Dallas, Texas. In June 2002, he initiated his research work as a Ph.D. candidate in the laboratory of Drs. John Minna in the Hamon Center of Oncology and David Mangelsdorf in the Howard Hughes Medical Institute and Department of Pharmacology. On March 20th, 2007, he defended this dissertation and expects to receive the Ph.D. degree from U.T. Southwestern in May, 2007.


Permanent Address:          535-9, Jeon-ha 1 dong, Dong-goo, Ulsan
                            682-061, Republic of Korea