12-2004

Analysis of Benzo(A)Pyrene-Induced Gene Expression Changes in Fisher 344 Rat Liver by Affymetrix Microarray and Quantitative Real-Time PCR

Alhaji Umar N'jai
Western Michigan University

ANALYSIS OF BENZO(A)PYRENE-INDUCED GENE EXPRESSION CHANGES IN FISHER 344 RAT LIVER BY AFFYMETRIX MICROARRAY AND QUANTITATIVE REAL-TIME PCR

by

Alhaji Umar N'jai

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Chemistry

Western Michigan University
Kalamazoo, Michigan
December 2004

NOTE TO USERS

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGMENTS

All thanks due to God almighty for giving me the strength and good health to complete this project.

I am extremely grateful to Dr. Jay C. Means for his constant support, encouragement, and for giving me the independence to carry out the research work and learn from my mistakes. Special thanks to all my committee members for their suggestions and inputs in the course of this research. Thanks also to all members of the WMU GLEAMS Center for their collaboration in the initial period of this project.

I enjoyed the positive and familiar environment of the Chemistry Department. Thanks to all the faculty members and staff from the chemistry department and the graduate college.

My sincere and wholehearted gratitude goes to my entire family and friends for their encouragement, inspiration and support. Special thanks to Papa and Dior. Finally, this dissertation is dedicated to my Late Dad and my mum for the love, support and the wonderful gift of education they gave to me.

Alhaji Umar N'jai

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

TABLE OF CONTENTS - Continued

TABLE OF CONTENTS - Continued

TABLE OF CONTENTS - Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES - Continued

LIST OF FIGURES - Continued

LIST OF TABLES IN APPENDIX

LIST OF FIGURES IN APPENDIX

CHAPTER I

INTRODUCTION

Urban air pollution caused by combustion of fossil fuels, refuse burning, domestic heating systems, vehicle emissions, and coke ovens constitute a major environmental health problem and implicated in diseases such as cancer (U.S. EPA, 1988; Bjorseth and Ramdahl, 1985). For instance, cancer is an important public health concern in the United States and around the world. After heart disease, it is the second leading cause of death, accounting for 23.0% of deaths in the U.S (Simmonds, 2003). It is estimated that the number of invasive cancer cases in U.S. will double by the year 2050 (Simmonds, 2003). Among men, the most common cancers are cancers of the prostate (33%), lung and bronchus (14%), and colon and rectum (11%). The three most diagnosed cancers among women, on the other hand, are breast (32%), lung and bronchus (12%), colon and rectum (11%). The incidence of lung cancer among women and in the general population has been on rise and is expected to be the leading cause of cancer deaths in the coming years (Simmonds, 2003).

Benzo[a]pyrene (BaP), a five ring member of the polycyclic aromatic hydrocarbon (PAH) family present in urban air pollution and also a product of tobacco smoke, grilled/barbecued meat, smoked fish, or green vegetables, is a potent carcinogen implicated in various cancers, lung, cardiovascular and other diseases in both human and animal studies (Wei et al, 2002; Kazerouni

et al, 2001; Yan et al, 2000; Pei et al, 1999; IARC, 1983). The United States Environmental Protection Agency (U.S.EPA) classified BaP as a "possible human carcinogen" (Group B2) and the International Agency for Research on Cancer (IARC) considers BaP a known animal carcinogen and a probable human carcinogen (Group 2A).

BaP and its metabolites are capable of causing this numerous health effects and disease such as cancer, immunogenesis, atherosclerosis, and endocrine disruption because of their ability to bind DNA and form BaP-DNA adducts, causing gene mutations (Chiang, 2001). Also, BaP and its metabolites can cause cell cycle regulations, transcriptional regulation, immunosuppression, oxidative stress, and increased intracellular calcium in cells (Jeffy et al, 2000; Yan et al, 2000; Canova et al, 1998). For instance, germline mutations of *BRCA1* gene are found in half of breast-ovarian cancer pedigrees and approximately 10% of women with early onset or sporadic breast cancers (Harkin, 2000). In a study by Jeffy et al. (2000), treatment of MCF-7 breast cancer cells with BaP and its metabolite Benzo[a]pyrene diol epoxide (BPDE) caused disruption of *BRCA-1* expression and cell cycle kinetics in breast epithelial cells. Subsequent studies (Jeffy et al, 2002) show that BaP lowers *BRCA-1* expression through repression of the transcriptional activity of its promoter fragment, which triggers a cascade of non-mutational events that leads to the onset of sporadic breast cancers. Similarly, BaP is also associated with tissue vitamin A depletion (Edes et al, 1992), an impairment leading to skin and other epithelial cancers. This is supported by

epidemiological data that associates low vitamin A status with lung cancer in humans. It has been shown to induce hepatic tumors in several animal models, tissue types and cell systems including rats, mice, fish and humans (Binkova et al, 2000). In addition, BaP and its metabolites are also linked to endocrine disruption (Vinggaard et al, 2000; Tian et al, 1999) in humans and animals by way of their anti-estrogen and androgen receptor activity. BaP and other PAHs may act as antiestrogens by, 1) binding to the aryl hydrocarbon (Ah) receptor leading to the induction of Ah responsive genes that result in broad spectrum of antiestrogenic response, 2) by blocking activation of the estrogen receptor (ER) in transfected cells, and 3) by simple binding to ER of BaP or its hydroxylated metabolites. Inhibition of the action of the androgen receptor through binding of BaP and other PAHs during the embryonic stage often leads to alterations in the development of the male external genitalia.

Studies in rodents and humans indicate BaP and its metabolite are immunotoxic to T and B cell lymphocytes (Davila et al, 1995; Mounho and Burchiel, 1998; Ladics et al, 1992). BaP and its metabolite (BPDE) caused significant elevation of intracellular calcium levels in normal human peripheral blood mononuclear T and B cell (Mounho and Burchiel, 1998; Ladics et al, 1992). Elevation of intracellular calcium was found to trigger increase tyrosine phosphorylation, and tyrosine phosphorylation of two *src*-related protein tyrosine kinases, *lyn* and *syk*, which subsequently lead to immunosuppression. Apoptosis in Daudi human cells (Salas and Burchiel, 1998; Davila et al, 1995) was observed in response to BaP and BPDE

treatment. Apoptotic responses in immune cells are critical steps in many disease states including malignancy, neurodegenerative conditions, autoimmune diseases, and viral infections. Altered calcium signaling in apoptotic responses is emerging as a primary mechanism of toxic injury for lymphoid and non-lymphoid cells (Salas and Burchiel, 1998). Recent evidence suggests that, in addition to mutagenic effects, BaP may induce epigenetic events that contribute to the deregulation of liver growth and differentiation induced by the compound (Parrish et al, 1998). Epigenetic carcinogens such as peroxisome proliferators have been shown to increase the expression of protooncogenes in mouse epithelial cells following acute exposure, and the extent of gene induction was related to tumorigenicity (Bral and Ramos, 1997). Bral and Ramos (1997) have shown that BaP activates the *c-Ha-ras* proto-oncogene via a transcriptional mechanism that may involve the aryl hydrocarbon (Ah) receptor complex.

The mechanisms through which BaP and other PAH compounds exert these biological effects are still relatively poorly defined. Binding to the Ah receptor and the subsequent activation of the metabolic enzymes (cytochrome p450) is considered to be the major pathway for BaP toxicity. The mechanism for these Ah receptor mediated pathophysiological responses is not well understood. The Unliganded Ah receptor is located in the cytosol associated with heat shock protein 90 (*hsp90*). Upon ligand binding, *hsp90* is released from the complex, and the receptor translocates into the nucleus and dimerizes with the aryl hydrocarbon receptor nucleus translocator (*Arnt*)

protein (Barouki and Morel, 2001; Schmidt and Bradfield, 1996). The heterodimer binds to the xenobiotic response elements (XREs) and alters expression of genes controlled by enhancer XREs. XREs, with the conserved core sequences "GCGTG", are found in the promoter regions of several genes involved in the metabolism of BaP, including cytochrome p450 (*Cyp1a1, Cyp1a2,* and *Cyp1b1*), and NAD(P)H-quinone oxidoreductase. Oxidation of BaP by these metabolic enzymes produces reactive intermediates that can bind to cellular macromolecules such as DNA and form DNA adduct (Miller and Miller, 1976) and cause initiation of cancers and adverse health effects in mammals.

Our goal in this project is to identify biomarkers or signature genes associated with BaP toxicity. Biomarkers are commonly used in risk assessment and are useful for determining exposure to specific contaminants. Biomarkers are not necessarily indicative of an overall toxicological effect of a chemical, but indicate that exposure to a contaminant or class of contaminants has occurred. However, under certain circumstances, biomarkers may correlate with adverse health effects. *In vitro* and *in vivo* exposure to BaP is associated with change in expression of genes involved in apoptosis, anti-apoptosis, detoxification, DNA repair, oncogenes and so on. However, *in vivo* genome-wide transcriptional responses to BaP exposure in mammalian species such as rat are not fully characterized. More specifically, no such studies have examined the dose-response relationships and temporal (sub-acute and sub-chronic) regulation of gene expression by BaP *in vivo*. In addition, low dose

exposure to other compounds alters the mRNA transcript profiles of different cell or tissue types. Whether or not such low dose responses can occur in response to BaP exposure *in vivo* is not well characterized in the literature. The specific hypothesis behind the proposed research is that exposure to a low, medium and high dose of BaP at different time points alters the mRNA transcript profiles of Fisher 344 rat liver and causes distinct patterns of gene expression. That hypothesis is based on the following observations. First, BaP is a known carcinogen and forms DNA adducts in rat liver. Upon metabolic activation, BaP and its metabolites are capable of binding to DNA forming BaP-DNA adducts, which may lead to cancer (Chiang, 2001, cavalieri and Rogan, 1985). This may lead to increase in expression of apoptotic, anti-apoptotic, detoxification and DNA repair genes. Second, increase in reactive oxygen species is associated with BaP metabolism in the cell. This increase in reactive oxygen species may lead to increase expression of antioxidant genes such as glutathione-s-transferases, glutathione peroxidases and other genes (Barouki and Morel, 2001; Flowers and Miles, 1991). Third, studies with other chemical carcinogens have shown non-linear dose response with low dose exposure showing significant increases in gene expression over the high dose exposures. The gene expression tends to differ sharply with the nature of the dose. For instance, genes involved in DNA repair and detoxification are highly expressed under high dose conditions. Fourth, temporal regulation is expected to play an important role in expression profile of various transcripts. For example, adaptive responses can occur in

response to a xenobiotic challenge under low dose conditions. The nature of the dose and the duration or extent of exposure may lead to distinct patterns of gene expression (Hossain et al, 2000; Yoneda et al, 2001).

Based on these observations, the experimental focus of this research project is to use DNA microarrays to evaluate the *in vivo* genome-wide transcriptional response to BaP challenge or insult in Fisher 344 rat liver. Toxicologists traditionally measure physical parameters (example, body weight, organ weight, blood pressure, and activity level), histological features of tissue samples (histopathology), or blood chemistry indicators to identify adverse effects of exposure. Such data are useful diagnostic indicators, but they are less useful for understanding the molecular mechanisms of toxicity or to gain insight into the pathogenesis and progression of disease. In addition, traditional toxicological methods are usually insufficiently sensitive to detect low-level toxicity or early pre-clinical stages of disease. Moreover, previous studies that defined molecular indicators of exposure and health effects were limited to assaying changes in a limited number of proteins or mRNAs.

With the advent of DNA microarrays, thousands of transcripts or genes can be assayed simultaneously. Another great advantage of microarray technology is the ability to determine relationships among genes and biological pathways that underlie subtle changes in physiology. Measurement of gene expression can potentially provide clues about the regulatory mechanisms, biochemical pathways and broader cellular function. In addition, the determination of genes expressed in disease tissue compared with their

normal counterpart helps in understanding the disease pathology and also in identifying potential points for therapeutic intervention. The rationale for measuring mRNA expression rather than the protein product is based on the efficiency of obtaining global and quantitative information about gene expression. Moreover, the RNA-based approaches to assess gene expression compared to proteomic technologies are considered to be more robust and reproducible.

Gene expression data obtained with microarrays need to be validated with Real-Time Quantitative Polymerase Chain Reaction (RT-QPCR) or other alternative methods. Real-time PCR quantitation of nucleic acids is based on detection of amplified products at the end of each cycle, which permits quantitative calculation of target RNA in an unknown sample by comparison with the kinetics of PCR product accumulation in samples of known quantity. Real-time PCR product accumulation is monitored during amplification by the hybridization of an additional dual-labeled gene-specific oligonucleotide probe that allows detection during PCR. Quantification of the transcription levels of these genes plays a central role in the understanding of gene function and of abnormal alterations in regulation that may result in a disease state.

Recently, epidemiologists have begun employing these early markers of biologic effect because of the challenges in using cancer as an outcome measure in occupational epidemiologic studies. The use of clinical disease as an endpoint in cases such as cancer is limited by the problem of latency; sometimes 10, 20, or even 30 years between exposure and disease. This has

spurred the growth of molecular toxicology and epidemiology to characterize new generation of biologic markers, with application to occupational diseases, and allow identification of excess risk in the natural course of a disease. This provides an opportunity for preventive action. Another benefit of early markers of disease include the ability to enhance exposure assessment, especially low-dose exposures and low risk populations, identify risks from single agents within complex exposures, estimate the total exposure from multiple sources, and provide data for predictive toxicological assessment.

CHAPTER II

LITERATURE REVIEW

Properties and Sources of Benzo(a)pyrene

Benzo(a)pyrene (BaP) is a five-ring polycyclic aromatic hydrocarbon (PAH), also known as 3, 4-benzyrene, which was first isolated from coal tar. It was later found to be a component of cigarette and marijuana smoke. BaP as a member of the PAH family, only comprises less than five percent of the total atmospheric amount of polycyclic aromatic hydrocarbons (PAHs). BaP has a molecular formula of $C_{20}H_{12}$. It is a yellowish crystal with a molecular weight of 252.3. The boiling point of BaP is greater than 360° C at 760 mm Hg and the melting point is approximately 179° C. The vapor pressure of BaP at 25° C is 8.4×10^{-7} pa, enthalpy of sublimation at 25° C is 118.3 ± 2.2 KJ/mol, entropy of sublimation at 25° C is 183.6 J/mol EK, enthalpy of fusion is 16.6 ± 0.3 KJ/mol, log P (octanol/water) is 6.04, and the Henry constant is 2.7×10^{-7} Atm-m^3/mol (Bjorseth, 1983). It is soluble in benzene, xylene, and toluene. It is sparingly soluble in ethanol, methanol, and insoluble in water.

PAHs and BaP are formed as products of the incomplete pyrolysis of organic materials and are present in considerable quantities in fossil fuel from which they are released by a variety of combustion processes. The principal sources of BaP and other PAHs in the environment therefore include, in the wider sense, power plants, domestic heating systems, petrol and diesel

engines, refuse burning, vehicle emissions, coke ovens, and various industrial activities, while tobacco smoke provides a more localized source of supply. Each of these sources of PAH produce a mixture containing between 100 and 300 different individual hydrocarbons, and the estimated total annual emission of BaP in the United States (US) is some 1200 tons (Kazerouni et al, 2000; Harvey, 1991). Furthermore, emissions from these sources are major contributors to urban environmental contamination, particularly in areas adjacent to highways, industries, and airports. Although, there are natural sources of PAHs (volcanic activity and forest fires, the anthropogenic sources are still considered to be the most significant sources of BaP and other PAHs in the environment.

PAH are also generated in foods through pyrolysis during charbroiling of meat products and through contamination of foodstuffs by contact with either petroleum products or coal tar products. Because of the geochemical processes and atmospheric deposition of air pollution particulate on the crops, it is possible to generate these naturally occurring PAHs in food. For instance, Kazerouni et al (2001) observed higher BaP content in green leafy vegetables such as kale and collard greens, compared to other vegetables in their study, which could be explained by their greater contact surface to the ambient air during growth and their late harvest during the winter months. A similar Italian study (Lodovici et al, 1995) found high levels of carcinogenic PAHs in all vegetables except tomatoes exposed to polluted air. Therefore,

the PAH levels in uncooked foods largely depend on the origin of the food and subject to regional variations.

PAHs are widely distributed throughout the waters of the earth surface, entering the food chain by being taken up by fish and plankton (Harvey, 1991). The occurrence of BaP in seafood may be attributed to aquatic pollution as opposed to cooking methods, unless they were smoked. Fish, shell fish, and other seafood may be contaminated by petroleum pollution indirectly through water contact in the marine environment. Since some PAHs including BaP are only slowly degraded, they represent a potential health hazard to humans through drinking water and other sources. For example, the average level of BaP in drinking water is estimated to be approximately 0.01μg/L, which is of a similar level to that from breathing reasonably clean air (Harvey, 1991)

Routes of BaP Entry

The most common way BaP enters the body is through the lungs when a person breathes in air or smoke. It may also enter the body through the digestive system when substances containing it are swallowed. Dermal exposure through the skin can also occur through contacts with soil (hazardous waste site) that contains high levels of BaP or through contacts with heavy oils that contain the compound (Agency for Toxic Substances and Disease Registry, 1990). BaP is present in a wide variety of food items (IARC, 1983). PAH are found in grilled/barbecued meat, vegetables, oils,

grains/cereals, fruits, smoked fish and seafood in concentrations as low as 0.001 ng/per g (200 ppb) (IARC, 1983). In human epidemiological studies, Kazerouni et al (2001) analyzed 200 food items for BaP and estimation of BaP intake. The highest levels of BaP (up to about 4ng BaP/g of cooked meat) were found in grilled/barbecued very well done steaks and hamburgers and in grilled/barbecued well done chicken with skin. BaP levels were lower in meats grilled/barbecued to medium done and in all broiled or pan-fried meat samples regardless of doneness level (Kazerouni et al, 2001). The Total Human Exposure to Environmental Substances (THEES) study has shown that people in the general population without substantial exposure to pollution and occupational exposure had a higher exposure to carcinogenic BaP by food ingestion than by inhalation (Butler et al., 1993). This exposure pathway is consistent with a large pool of epidemiological data linking the incidence of human cancer (20-50%) with dietary factors.

Many PAHs are potent carcinogens in animal models of chemical carcinogenesis, and the ability of PAHs to compromise the immune system is well documented in both *in* vitro and *in vivo* systems. Humans are exposed concurrently or sequentially to complex mixtures containing carcinogenic PAHs up to several thousand micrograms per day depending on the source, lifestyles, such as occupation, environmental food chain, cigarette smoking habits, and socioeconomic status (Arif and Gupta, 1996). However, there are no comprehensive PAH database, therefore, epidemiological studies have not been able to estimate dietary PAH intake and investigate the risk of cancer

and other diseases associated with it. However, various investigators (Kazerouni et al, 2001; Ward et al, 1997) have evaluated the relationship between the intake of several foods that may have high level of PAH (example smoked or grilled/barbecued meats) and risk for cancer at several sites including stomach and esophagus, colorectal, pancreatic and bladder cancer. Because PAH are contained in wide variety of foods, it is necessary to directly estimate the dietary intake of BaP from all dietary sources to evaluate the relationship between dietary intake of BaP and risk of cancer.

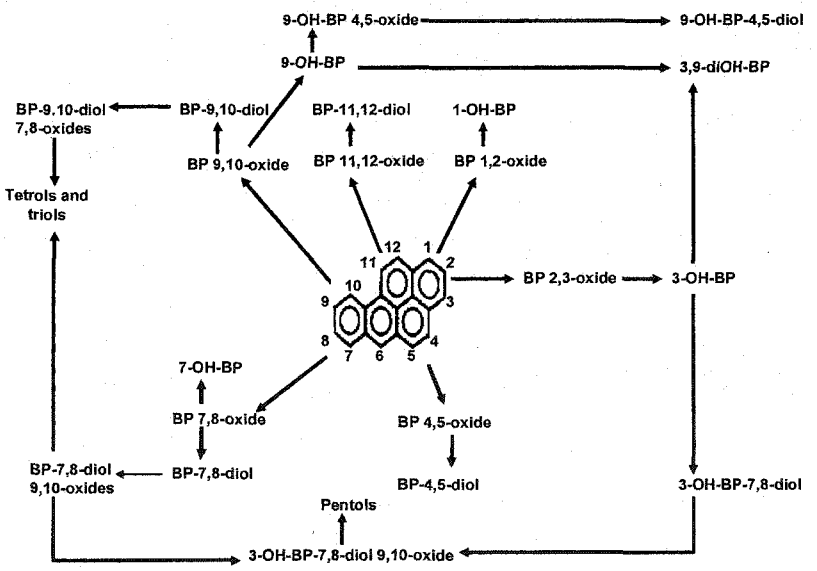


Figure 1: Pathways in the Metabolism of BaP to Hydroxylated Metabolites.

Health Effects of BaP

Exposure to PAH-contaminated air pollutants has been associated with the occurrence of cancer and pulmonary diseases. BaP found in cigarette smoke and air pollutants, is one of the most widely studied and potent PAH

carcinogen in animal experiments. A study by Weyand et al. (1995) showed an association between the ingestion of a diet containing BaP and forestomach tumors in A/J strain of mice. Another study by Culp et al. (1998) examined the histopathological sections of other sites such as small intestine, tongue and esophagus in female B6C3F1 mice. This study showed that BaP treatment in a 2-year feeding study resulted in an increased incidence of papillomas and or carcinoma of forestomach, esophagus and tongue.

The United States Environmental Protection Agency (U.S. EPA) classified BaP as a "possible human carcinogen" (Group B2) and the International Agency for Research on Cancer (IARC) considers BaP a known animal carcinogen and a probable human carcinogen (Group 2A). The Office of Environmental Health Hazard Assessment (OEHHA) agrees with the IARC Group 2A classification of BaP as a probable human carcinogen based on sufficient evidence for carcinogenicity in animals and limited evidence in humans. BaP has the ability through its metabolite to arylate DNA, causing gene mutations in prokaryotic and eukaryotic cells, induce sister chromatin exchanges in mammalian cells, and produce unscheduled DNA synthesis in mammalian cells (U.S. EPA, 1998).

Several types of malignant tumors have been induced in rodents by BaP. Tumor formation has been induced in a number of tissues, such as lung, skin, forestomach, liver, the hematopoieitic system, and the mammary gland (IARC, 1983). The National Institute for Occupational Safety and Health (NIOSH) has determined that workplace exposure to coal products can

increase the risk of lung and skin cancer in workers. Based on this, NIOSH suggests that the workplace exposure limit for coal tar products is 0.1 milligram of PAHs per cubic meter of air (0.1 mg/m^3) for a 10-hour workday and 40-hour work week. Although NIOSH has not suggested a specific workplace limit for BaP, the Occupational Safety and Health Administration (OSHA) has set a legal limit of 0.2 milligrams of all PAHs per cubic meter of air (0.2 mg/m^3).

Similarly, cardiovascular disease remains one of the major causes of death in industrialized nations and many have suggested that tobacco smoking is a major contributor to cardiovascular disease. Smoking promotes platelet hyperreactivity, accelerates atherosclerotic plaque formation, and increases ischemic tissue damage. Experimental evidence (Yan et al, 2000) suggests that BaP accelerates smooth muscle proliferation and promotes atherosclerosis in animals ranging from chicken to rats. In addition, Yan and others (2000) have also detected BaP-related DNA adducts in atherosclerotic arteries.

Carcinogenic PAH such as BaP, 7-12-dimethylbenz (a) anthracene (DMBA), and 3-methyl-cholanthrene (3-MC) are immunotoxic to rodent and human lymphocytes (Davila et al, 1995; Mounho and Burchiel, 1997). Both BaP and DMBA metabolites have been shown to suppress the *in vitro* plaque-forming cell (PFC) response to sheep red blood cells (Ladics et al., 1991). Humans exposed chronically to atmospheric complex mixtures of pollutants, composed primarily of PAHs, also develop suppression of their immune

system. For example, Szczeklik et al. (1994) examined the effect of PAH exposure on humoral immunity in a group of male iron foundry workers in Poland who were chronically exposed to PAHs, as assessed by personal and area monitoring. An evaluation of serum from these men revealed that coke oven workers exposed to high concentrations of atmospheric PAHs, including benz[a]anthracene (BA) and BaP, had reduced serum levels of immunoglobins *IgG* and *IgA* compared with mill workers exposed to concentrations of PAHs 3-5 magnitudes lower than the coke oven workers. Furthermore, the risk of cancer is greater in humans who are immuno-compromised. For instance, children with inborn deficiencies of the immune system have a high incidence of malignant tumors.

Also, a growing body of literature identifies PAHs as potential environmental endocrine disruptors (Vinggaard et al, 2000). Several pieces of evidence indicate that environmental chemicals, which are able to bind to the androgen receptor (AR), may have an important impact on abnormalities associated with developing male reproductive system. An inhibition of the action of the AR during the embryonic stage very often leads to alterations in the development of the male external genitalia such as cryptorchidism and hypospadia. Similarly, PAHs may act as antiestrogens by binding to the Ah receptor (Vinggaard et al, 2000) leading to induction of Ah responsive genes that result in a broad spectrum of antiestrogenic responses.

Chemical Carcinogenicity of BaP

Chemical carcinogens were first recognized in humans approximately two and half centuries ago by Hill in 1761, who noted an increase incidence of nasal polyps and cancer after the prolonged and excessive use of tobacco snuff. Also, the incidence of scrotal skin cancers in European chimney workers was correlated by Percival Pott to the sub-chronic or chronic exposure of these workers to coal soot and tar. There are three broad classes of chemical carcinogens; Direct acting (require no metabolic activation), Indirect acting (require metabolic activation), and Non-genotoxic or Epigenetic carcinogens (Waalkes and Ward, 1994). Majority of the chemical carcinogens including BaP are indirect acting, requiring metabolic activation to produce electrophilic intermediates in order for the tumor formation process to occur.

BaP after metabolic activation is known as one of the most potent PAHs to cause carcinogenic effects. It was first determined to be a carcinogen in the early 1950s by Elizabeth Miller who found that a metabolite of BaP reacts with proteins in mouse skin (Miller, 1951). BaP is potent because it has two main features: (1) a relatively low ionization potential of 7.23eV (IP), which allows metabolic removal of one electron forming a cation (Cavalieri and Rogan, 1985), and (2) appreciable charge localization in the radical cation, allowing the intermediate to efficiently react with nucleophiles (Josephy, 1997). This is evident in the most widely studied adduct formation of BaP with DNA and other macromolecules. A review of the literature on DNA adducts formation of BaP shows that it is activated by two main routes, 1) one-electron oxidation

to yield reactive intermediate radical cations (cavalieri and Rogan, 1985), and 2) monoxygenation to produce bay- region diol epoxides (Cheh et al, 1993). In the former, the major identified BaP diol epoxides adduct arises from formation of a covalent bond between C-10 of BaP-diol epoxide and two amino of guanine moieties on DNA. In the latter, on the other hand, BaP is first activated by one-electron to its radical cation, and binds the C-6 of BaP to either the C-8 or N-7 of guanine on DNA and forms adducts (Rogan et al, 1990). In both cases, the activated intermediates react with nucleotides to form covalently bonded compounds, causing "molecular lesions."

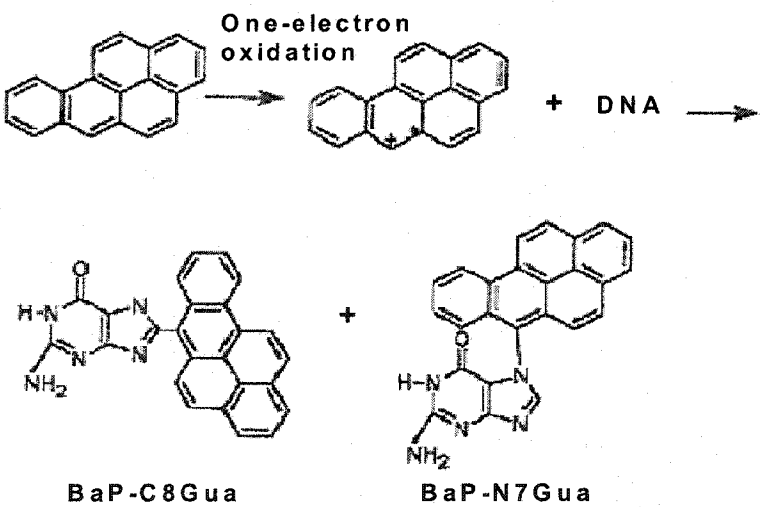


Figure 2. Formation of BaP-DNA Adducts by One-electron Oxidation

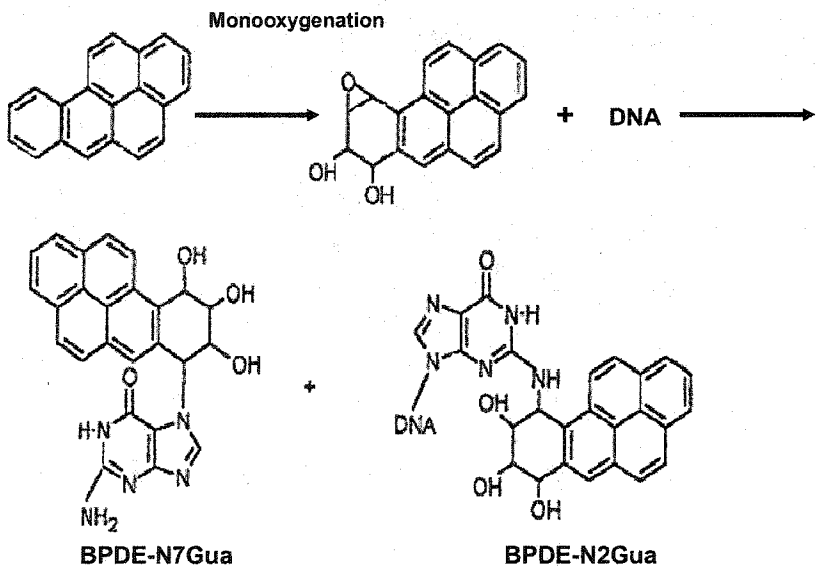


Figure 3. Formation of BaP-DNA Adducts by Monoxygenation

Molecular Mechanisms of BaP Toxicity

The mechanisms through which PAH compounds exert their biological toxic effects are still poorly defined. Binding of this class of compounds to the aryl hydrocarbon (Ah) receptor, which then activates the expression of metabolic enzymes such as cytochrome p450 1a1 (*Cyp1a1*) and 1a2 (*Cyp1a2*), is acknowledged to be a critical step in the carcinogenicity or immunotoxicity of these compounds. The process of transcriptional activation of these cytochromes is felt to proceed in large part through a signaling process involving ligand- dependent activation of the Ah receptor. The cytosolic Ah receptor belongs to the family of the basic helix-loop-helix (bHLH) proteins and is occupied in the cytoplasm of the cell with heat shock protein of 90 kDa

(*hsp90*), which is uncoupled from the receptor in the presence of ligand (Chen and Tukey, 1996; Denison and Heath-Pagliuso, 1998; Schmidt and Bradfield, 1996). Following Ligand binding, the Ah receptor is released from a cytosolic multiprotein complex (at least two hsp90 and an additional 37 kDa Ah receptor interacting protein) to which it is associated. The liganded Ah receptor complexes subsequently translocate into the nucleus, where it associates with the bHLH Ah receptor nuclear translocator (Arnt) protein, and possibly other factors to form the high affinity DNA binding Ah-Arnt heterodimer. The Ah-Arnt heterodimer complex then binds to specific enhancer sequences, the dioxin response element (DRE), upstream of the *Cyp1a1* gene (as well as other responsive genes), which leads to DNA bending, chromatin and nucleosome disruption, increased promoter accessibility and increased *Cyp1a1* gene transcription (Denison and Heath-Pagliuso, 1998). The Ah-Arnt complex synergizes with other transcription factors such as *NFI* and *SP1*, thereby activating the transcription of the gene. Several other genes are also the target of the Ah receptor. These include other Cyp450 genes such as *Cyp1a2, Cyp1b1*, and several phase II genes such as *NQO-1* and Glutathione-s-transferase (*Gst*) as well as a number of other genes.

Ligands of the Ah receptor, such as BaP and the dioxin tetrachlorodibenzodioxin (TCDD) trigger a number of toxic effects in rodents including cancer, immunosuppression, and endocrine disruption. In addition to its role in modulating the induction of gene expression, numerous studies

also support a role for the Ah receptor in mediating PAH toxicity. For instance, structure-activity relationships studies using a variety of inbred strains of mice which differ in their Ah receptor functionality reveals that the ability of PAHs and other xenobiotic compounds to bind to the Ah receptor not only correlates well with their ability to induce gene expression but also their ability to produce toxicity.

The ability of the Ah receptor to produce toxicity in cells is exemplified by the metabolism of BaP. BaP metabolism in the cell occurs through phase I and Phase II enzymatic oxidation and hydrolysis. Although the physiological function of the *Cyp1a1* (Phase I enzyme) seems to be detoxification, convincing evidence suggests that high activity of the enzyme could lead to toxicity. This is largely because the detoxification of the highly mutagenic BaP diol-epoxides as well as other reactive intermediates depends on the activity of phase II enzymes such as glutathione-s-transferase (*Gst*). Therefore, the ratio between *Cyp1a1* and phase II enzymes is critical to avoid the accumulation of putatively toxic intermediates of BP metabolism.

Another mechanism through which *Cyp1a1* could lead to toxicity is the production of reactive oxygen species (ROS). Several monooxygenases have been shown to release hydrogen peroxide (H_2O_2) during their catalytic cycle (Barouki and Morel, 2001). Experiments in cultured cells have shown that the induction of *Cyp1a1* by BaP, as well as the transfection of a *Cyp1a1* expression vector, leads to oxidative stress within the cells (Barouki and Morel, 2001). High *Cyp1a1* enzyme activity within the cell may be deleterious

because of the generation of an intracellular oxidative stress and the subsequent oxidation of biological macromolecules. Other studies (Flowers and Miles, 1991) have found alterations of pulmonary BaP metabolism by reactive oxygen metabolites. The activity of both BaP hydroxylase and the monooxygenases (cytochrome-p450 and NADPH cytochrome p-450 reductase) was inhibited by addition of reactive oxygen metabolites to rat lung microsomes. However, the reactive oxygen metabolites (superoxide anion and H_2O_2) had no effect on the formation of minor metabolites of BaP; the BaP-quinones and BaP-dihydrodiols (Flowers and Miles, 1991). This are the BaP metabolites thought to produce toxic effects and which may lead to the formation of carcinogens or mutagens. The results of these experiments suggest that exposure of lung microsomes to oxygen metabolites can lead to a slowing of overall BaP metabolism and the increased accumulation of potentially toxic BaP metabolites. The toxic consequences of high *Cyp1a1* activity correlate with epidemiological observations in humans; some studies have documented a correlation between high *Cyp1a1* activity and lung cancer.

Although *Cyp1a1* is induced primarily by Ah receptor activation, it is unlikely that this represents the only determinant of the level of expression of this gene. Several other factors such as hormones, cytokines, and chemicals have been shown to modulate this effect. The efficiency of the Ah receptor could be modulated by phosphorylation events. Other studies have established that a number of cytokines (tumor necrosis factor-α, *TNF-α*,

Interleukin-1β, *IL-1β*) also repress the induction of the Ah receptor gene. In addition, an autoregulatory loop control of *Cyp1a1* enzyme on expression of the *Cyp1a1* gene has been observed in experiments (Barouki and Morel, 2001). Barouki and Morel (2001) used cotransfection experiments to show that the over expression of *Cyp1a1* enzyme leads to the down regulation of the *Cyp1a1* gene promoter by targeting the same transcription factor, *NFI/CTF*. Also, other regulatory loops could also be involved in the fine tuning of *Cyp1a1* gene induction. For example, the Ah receptor repressor (AhRR), which is structurally similar to Ah receptor but lacks transcriptional activity, antagonizes the Ah receptor by competing for Arnt.

In addition, binding of PAHs, such as BaP, to the Ah receptor, which then activates the expression of metabolic enzymes, such as *Cyp1a1* and *Cyp1a2* is acknowledged to be a critical step in the immunotoxicity of these compounds (Davila et al, 1995; Salas and Burchiel, 1998; Mounho et al, 1997). However, certain PAHs such as 7, 12-dimethylbenzanthracene (DMBA) suppress both the humoral and cell-mediated immune function in Ah-non-responsive and responsive mice, suggesting that DMBA may bypass cytochrome p450 mediated metabolism. The immunosuppressive effects of PAH appear to be mediated through disruption of one or more cellular signal transduction pathways either through altered cytokine production and/or disruption of cell surface membrane signaling. Both DMBA and BaP have been shown to suppress T-cell *IL-2* secretion and expression in both *in vitro* and *in vivo* studies (Burchiel et al, 1987; Salas and Burchiel, 1998). Also, the

secretion of *IL-1* after lipopolysaccharide (LPS) stimulation was increased in macrophage exposed to DMBA, while BaP had no effect on IL-1 secretion. Most PAHs are also known to have effects on cell viability and cell proliferation in culture (Davila et al, 1995; Salas and Burchiel, 1998).

The mechanism(s) by which PAHs exert their immunotoxic effects have not been well characterized in human lymphocytes (Mounho et al, 1997; Davila et al, 1995). A change in intracellular calcium and induction of apoptosis in cells is associated with the signal transduction mechanism involved in lymphocyte toxicity (Davila et al, 1995; Salas and Burchiel, 1998). For instance, one of the immediate consequences of mitogen contact with responsive cells is the activation of various ion transport systems in the plasma membrane and changes in intracellular ionic composition (Salas and Burchiel, 1998). Calcium serves as a second messenger and therefore mitogen-induced alterations in the levels of cytoplasmic free calcium triggers and regulates cell proliferation. The mobilization of intracellular calcium leads to the activation of the signal transduction pathway. PAHs that are neither carcinogenic nor immunosuppressive cause only a transient rise in intracellular calcium in human T-cell lines, while potent carcinogens and immunosuppressant also produce a sustained elevation in intracellular calcium (Davila et al, 1995). The rapid rise in intracellular calcium in response to extracellular stimuli is mediated by the second messenger inositol 1, 4, 5-triphosphate, which triggers the release of calcium from the endoplasmic reticulum. Furthermore, Salas and Burchiel (1998) have demonstrated that PAHs, such

as DMBA and BaP, induce cell death by apoptosis in murine B and T lymphocytes via Ca^{2+}-dependent mechanisms. Apoptotic responses in immune cells are critical steps in many disease states including malignancy, neurodegenerative conditions, autoimmune diseases, and viral infections. Ca^{2+}-dependent endonucleases mediate DNA fragmentation, and Ca^{2+}-dependent pathways have been implicated in mechanistic studies of apoptosis in immune cells. The pathways for the apoptotic response are thought to involve Ca^{2+}-dependent protein kinases and phosphatases that exert transcriptional control over apoptosis regulatory molecules. There is also accumulating evidence that Ca^{2+} regulation is a primary target of gene products, such as *Bcl-2* gene family members that serve as molecular integration sites for signals that regulate mitogenesis, differentiation, and apoptosis (Salas and Burchiel, 1998).

Additionally, the phosphorylation of proteins provides for direct signal amplification and is an important way by which the activities of proteins can be modulated. Many products of oncogenes are tyrosine kinases and phosphorylation of these tyrosine residues causes conformation changes in the protein, allowing other proteins that contain specific tyrosine phosphate binding-sites to bind to the phosphorylated protein and activate other enzymes. For example, the *ras* oncogenes product, which is induced by BaP, is a membrane associated GTP-binding/GTPase and is a key component of signal transduction pathways, linking receptor tyrosine kinases with cytosolic signaling proteins that ultimately impinge on the cell nucleus. Activated *ras*

oncogenes have been identified in various forms of human cancers, most notably colon and pancreatic cancers (Marshall, 1995). Some oncogenes may be directly activated by interactions between the target genes and chemical carcinogens, while other oncogenes may be indirectly activated during the progression of chemical carcinogenesis by non-target genetic events (Brown et al, 1990). Studies (Brown et al, 1990) of mouse skin tumors initiated by different chemical carcinogens showed that for each chemical initiating agent, a distinct spectrum of *ras* mutations was observed and that the distribution of *ras* gene mutations differed between benign papillomas and carcinomas, which suggests that the nature of the molecular initiating event influences the probability of papillomas progression to malignancy.

Research Goals and Instrumentation

Our research goal is to identify and characterize molecular markers or biomarkers associated with exposure to BaP, in order to understand the relationship between environmental exposure and human disease susceptibility. In addition, we hope to characterize potential mechanisms of action of environmental contaminants through the identification of gene expression networks, which in turn will be useful to elucidate the biological processes that govern or underlie specific health effects such as cancer or immunogenesis. An integrated functional genomics approach is employed in this study to characterize the nature of the gene response associated with the *in vivo* exposure of Fisher 344 rats to BaP. The genomics techniques that

would be utilized are high density oligonucleotide microarrays (Affymetrix GeneChips, Affymetrix Inc, CA) and Real-Time Quantitative Reverse Transcriptase Polymerase Chain Reactions (RT-QPCR) to analyze the dose-response as well as temporal regulations associated with BaP-induced gene expression patterns in Fisher 344 rat livers.

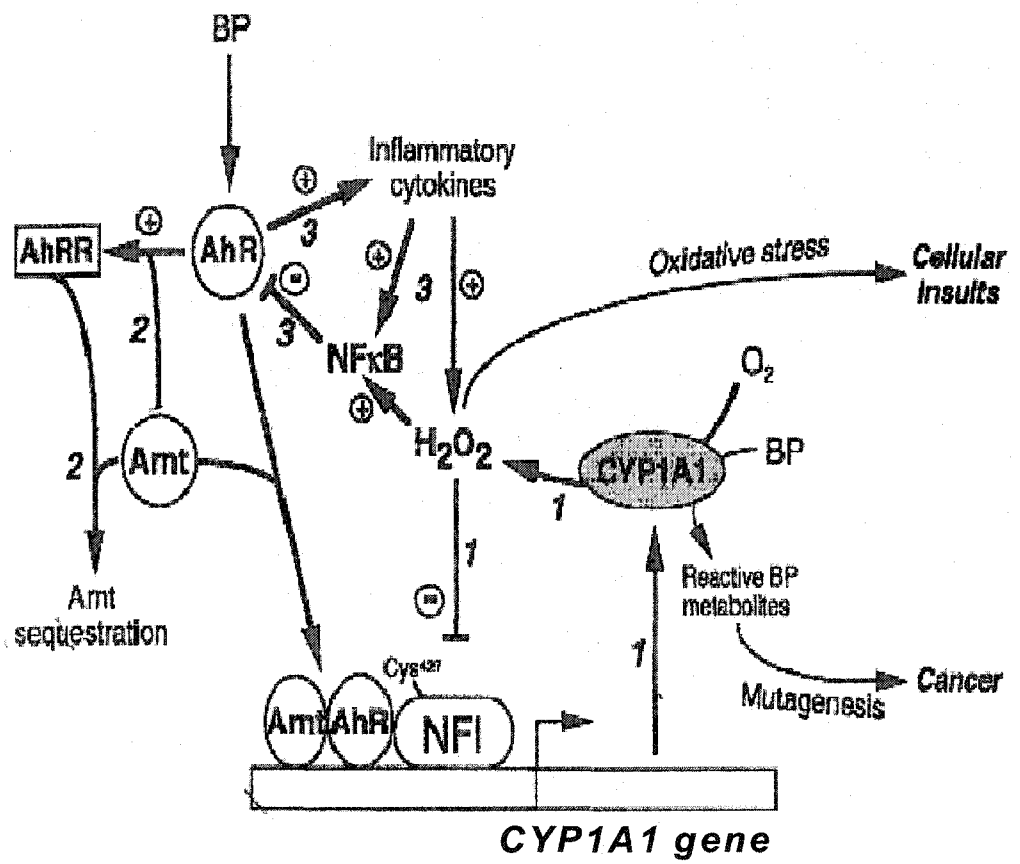


Figure 4. Possible Mechanism(s) of BaP Activation through the Ah Receptor Pathway.

Gene Expression Profiling and DNA Microarrays

Each year as hundreds of new drugs or chemical entities are introduced into the market or environment, a substantial challenge has been to determine the potential adverse health effects associated with exposure to such substances (Neumann and Galvez, 2002; Bartosiewicz et al, 2001). Risk assessment depends on the use of biomarkers or biological indicators of exposure to such compounds. Biomarkers are often used as a measure of exposure to a xenobiotic, and are physiological markers altered in response to chemical exposure. Biomarkers are not necessarily indicative of an overall toxicological effect of a chemical, but indicate that exposure to a contaminant or class of contaminants has occurred. However, under certain circumstances, biomarkers may correlate with adverse health effects. Traditional risk assessments to define molecular indicators of exposure and health effects were limited to assaying changes in a limited number of proteins or mRNAs. Currently, toxicological screening of these chemicals relies heavily on the use of a combination of short-term assays that often require separate bioassays to determine immunotoxicity, neurotoxicity, and genotoxic end points. These assays, however, are costly, require many animals, and may take several years and millions of dollars to complete. For example, the USEPA uses three-species bioassays to determine the relative toxicity in water of specific chemicals or specific effluents to wildlife. These tests are limited to lethality, growth, and reproductive end points, therefore, optimization of these processes is required in order to develop methodologies that are more

efficient, informative and cost effective (Bartosiewicz et al, 2001, Thomas et al, 2001).

With the advent of DNA microarrays, thousands of transcripts or genes can be assayed simultaneously. DNA microarrays are a powerful, high throughput tool for monitoring the expression of thousands of genes simultaneously. Considerable attention has been given to the application and use of DNA microarrays in drug discovery and development, in identifying changes in gene expression associated with various disease processes, and in screening populations for allelic variants or single nucleotide polymorphisms (Reninger et al, 2000; Afshari et al; 1999).

Public access to genome sequence data, accompanied by rapid advances in sequencing technology and molecular biology tools development, is enhancing the elucidation of the molecular mechanism of cell metabolism. Microarrays are instrumental in providing a holistic picture of the underlying genetic complexities that orchestrate biochemical pathways and networks. Moreover, microarrays are also been used to assess the expression of entire organism genomes, thereby paving the way towards global mapping and modeling of the regulatory mechanisms of genetic networks.

Another advantage of DNA microarrays is that they combine the ability to analyze the expression of a large number of genes with parallel data acquisition. This parallel-processing power allows experimental designs that are much less costly, time-consuming, and produce significantly more data

than conventional molecular methods, yet produce similar results (Bartosiewicz et al, 2001; Neumann and Galvez, 2002).

Recently, there has been considerable interest in using DNA microarrays for assessing the toxic effects induced by drugs and environmental contaminants on host physiology (Bartosiewicz et al, 2001; Waring et al, 2001; Afshari et al, 1999; Amin et al, 2004). Microarrays can be used to quickly classify toxicants based on characteristic expression profiles and to use these profiles as a means of identifying putative mechanisms of action (Amin et al, 2004), as well as for identifying genes that are co-expressed or under similar regulatory control (Waring et al, 2001). For example, the expression profiles associated with a heavy metal, a PAH, and an oxidant stressor may differ considerably and could be useful in assessing the potential toxicity of the new chemical entities, as well as, in environmental risk assessment monitoring.

DNA Microarrays Technology

DNA microarrays are rapidly becoming a fundamental tool in genomic research. DNA microarrays are an organized arrangement of multiple DNA probes fixed on an immobilized surface. DNA microarrays function using the classical hybridization techniques (Southern and Northern blots) of molecular biology. The key to the method is the property of nucleic acids to bind specifically and strongly to each other using Watson and Crick base pairing (cytosine to guanine, and thymine to adenine). The specificity of the binding is such that mismatches at a single position in a DNA strand of, for example,

20 bases, under suitable conditions sufficiently reduce the binding affinity as to completely prevent hybridization.

There are several microarray technologies, however, two approaches are currently prevalent; cDNA arrays and oligonucleotide arrays. Although they both utilize hybridization, they differ in how DNA sequences are laid on the array and in the length of these sequences. A detailed review of the technical aspects of these technologies is provided by Schena (2000). Briefly, in spotted DNA arrays, mRNA from two different biological samples is reverse-transcribed into cDNA, labeled with dyes of different colors, and hybridized to DNA sequences, each of which is spotted on a small region, or spot, on a glass slide. Spotted arrays are manufactured using x-y-z robots that use hollow pins to deposit cDNA, clones, or short and long oligonucleotides onto specially coated glass microscope slides (Schena et al, 1995). Spot sizes range between 50 and 500 μm in diameter depending on the type of pin used. After hybridization, a laser scanner measures dye fluorescence of each color at a fine grid of pixels. Higher fluorescence indicates higher amounts of hybridized cDNA, which in turn indicates higher gene expression in the sample. The strength of this technology is its flexibility and also allows the simultaneous interrogation of two biological samples such as untreated cells versus treated cells, or healthy tissue compared with cancer. The major disadvantage of spotted arrays is due to the variability in spot quality from slide to slide, and in addition, they provide information only on the relative

gene expression between specific cells or tissue samples as opposed to direct quantification of RNA expression.

High density oligonucleotide GeneChips (Affymetrix, Santa Clara, CA), on the other hand, are produced by synthesizing tens of thousands of short oligonucleotides in situ onto glass wafers, one nucleotide at a time, using a combination of photolithography and light-directed solid-phase DNA synthesis (Lipshultz et al, 1999).

In Affymetrix arrays, expression of each gene is measured by comparing hybridization of the sample mRNA to a set of probes, composed of 11-20 pairs of oligonucleotides, each of length 25 base pairs. The first type of probe in each pair is known as perfect match (PM) and is taken from the gene sequence. The second type is known as the mismatch (MM) and is created by changing the middle (13^{th}) base of the PM sequence to reduce the rate of specific binding of mRNA for that gene. The MM serve to control experimental variation and nonspecific binding of mRNA from other parts of the gene. A detail description of Affymetrix GeneChips technology is discussed by Lockhart et al. (1996). Affymetrix GeneChips are ideally suited for large scale gene expression analysis, where high quality is desired and the same array design can be used repeatedly.

Another advantage is the ability to measure the absolute expression of genes in cells or tissues. The sensitivity of the GeneChip enables detection of mRNAs present at levels as low as 1 transcript in 100000 (Lipshultz et al, 1999).

DNA Microarrays and Toxicogenomics

It is postulated that measurement of gene expression can potentially provide clues about the regulatory mechanisms, biochemical pathways, and broader cellular function associated with exposure to a xenobiotic compound (Afshari et al, 1999). Gene transcription lies at the beginning of the transcription of a cell to a xenobiotic. Therefore a transcriptional response can give a preliminary indication of the biochemical or biological mechanism being affected by a xenobiotic, and gene expression data can provide starting points in a toxicological examination. The rationale for measuring mRNA expression rather than the protein product is based on the efficiency of obtaining global and quantitative information about gene expression. Moreover, the RNA-based approaches to assess gene expression compared to proteomic technologies are considered to be more robust and reproducible. Toxicogenomics attempts to define how the regulation and expression of genes mediate the toxicological effects associated with exposure to a chemical. The basic premise is that a certain degree of similarity or generalized pattern of gene expression is expected during exposure to different classes of toxicants (Afshari et al, 1999). Also, subtle differences in gene expression patterns induced by chemicals belonging to a specific class of toxicants may be sufficiently distinct so as to identify chemical-specific signatures of exposure. Current toxicogenomics studies have focused on classifying toxicity according to unique gene transcriptional profiles induced by different classes of compounds.

According to Afshari et al (1999), gene expression profiling using DNA microarrays may be used to 1) Characterize potential mechanisms of action of environmental contaminants through the identification of gene expression networks, 2) Identify modes of action for previously uncharacterized toxicants based on correlations with the molecular signatures of toxicants with well characterized etiologies, 3) Assess toxicant-induced gene expression as a biomarker of chemical exposure, 4) Ascertain overlapping patterns of gene expression in animal species exposed to specific toxicants, 5) Use microarray information to extrapolate effects of toxicants from one animal species to another, 6) Characterize the biological effects of complex chemical mixtures, 7) Examine the effects of chronic versus acute exposure to chemicals, and 8) characterize genetic polymorphisms in populations and assess the importance of genetic polymorphisms in modifying individual susceptibilities to a contaminant or toxicant.

In spite of the wide range of possible applications of DNA microarrays, several assumptions are inherent to a toxicogenomic-based approach in toxicology, many of which remain to be validated. Some of these assumptions relate to the following: First, the influence of toxicant exposure on convergence and divergence patterns of gene transcription. It is generally assumed that gene transcription profiles induced by different contaminants may either converge to produce similar toxicological effects or similar gene expression profiles may eventually diverge at later time points to produce different toxicological effects. For example, although the toxicity of PAHs

such as BaP is related to Ah receptor binding and subsequent induction of cytochrome p450 genes (Bartosiewicz et al, 2001; Mounho et al, 1997; Davila et al, 1995), sometimes binding to intracellular signaling proteins downstream of the Ah receptor may affect a multitude of distinct biochemical pathways. In some cases, genes with apparently opposite effects may be induced, even under the same treatment conditions, as shown by Burczynski et al (2000), who observed induction of both apoptotic and anti-apoptotic genes in a human cell line in response to treatment with cisplatin. On the other hand, gene expression profiles induced by different contaminants as observed by Waring et al (2001) may converge at distinct temporal or spatial points to yield similar toxicological effects. For instance, Waring et al (2000) observed that treatments of rats with methapyrilene (250 mg/kg/day for 3 days) induced gene expression patterns in liver tissues that were more closely related to arsenic exposure than to other related hepatotoxins.

Second, the influence of temporal and spatial parameters on gene transcription is an important assumption in any toxicogenomic study. In classical toxicological studies, acute and chronic dose responses are utilized as distinct endpoint measurements. Toxicity, on the other hand, represents a continuum of possible effects governed by both temporal and spatial factors that are contingent upon exposure conditions. Therefore, the perceived toxicological properties of any chemical are dependent on the route, dose, and duration of the exposure, and as such gene expression patterns are also subject to these variables (Neumann and Galvez, 2002) The data obtained

using DNA microarrays gene expression profiling represent a snapshot of the gene transcriptional responses occurring at a particular time (temporally restricted) and within a particular tissue (spatially restricted). Gene transcriptional programs in biological systems are intricately controlled by temporal (early, mid, and late-induced genes) and spatial mechanisms such as alternative RNA splicing. Hence, correct interpretation of DNA microarray data for the assessment of the toxicological properties of chemicals will require that temporal and spatial gene expression profiles be accounted for. Finally, it is also postulated in the literature that clustering of gene expression responses among different chemicals may be predictive of the mechanisms of toxicity (Waring et al, 2000; Afshari et al, 1999; Thomas et al, 2002). However, as shown in experiments by various researchers (Thomas et al, 2001; Waring et al, 2000), the toxicity data may not necessarily equate to a common mechanism of action even for similar classes of chemicals. This also depends on the type of clustering algorithm used to group chemicals into various toxicological classes.

Applications of DNA Microarrays in Toxicogenomics Studies

Recently, a number of studies (Waring et al, 2001; Afshari et al, 1999; Thomas et al, 2002; Burton et al, 2002; Bartosiewicz et al, 2001; Amin et al, 2004) have been published which illustrate the use of DNA microarrays for toxicogenomics analysis. Using microarray gene expression profiling, Hossain et al. (2000) were able to identify a potential mode of action for lead-induced

neurotoxicity characterized by a loss of integrity of the blood-brain barrier. Exposure of human fetal astrocytes to lead in culture resulted in a 3-fold increase in mRNA expression of vascular endothelial growth factor (Hossain et al, 2000). Furthermore, gene expression profiling revealed that lead exposure induced the expression of a variety of stress response genes, amino acid biosynthesis genes, and transfer RNA synthase genes in astrocytes, an observation that correlated with other known mechanisms of action of lead on these cellular processes.

Using custom arrays with a relatively limited gene set, Bartosiewicz et al. (2001) has examined the gene expression patterns in liver and kidney in response to five classes of chemicals (PAH: benzo(a)pyrene, 3-methylcholantrene; DNA alkylators: dimethylnitrosamine, ethylnitrosurea; peroxisome proliferators: diethylhexypthalate, clofibrate; heavy metals: $CdCl_2$, $HgCl_2$; and oxidative stressors: CCl_4, bromobenzene). Each pair of chemicals yielded a similar pattern of gene expression distinct from the other four classes of chemicals in time course and dose-response studies in mice (Bartosiewicz et al, 2001). Moreover, for each of the five classes of compounds tested, several induced or repressed genes were in common in each chemical exposure, whereas other genes were unique for that specific class of compound (Bartosiewicz et al, 2001). Similarly, Waring et al. (2001) compared rat hepatocytes gene expression profiles among 15 well-characterized hepatotoxicants. A strong correlation between known mechanisms of hepatoxicity and clustering of specific gene sets were

observed, indicating that chemicals with similar toxicological effects elicit closely related patterns of gene expression. For example, using hierarchical cluster analysis, aroclor 1254 and 3-methylcholanthrene elicited similar patterns of gene expression in microarrays (Waring et al, 2001). In other studies, DNA expression microarrays have been employed to characterize temporal regulation of gene expression (Yoneda et al, 2001; Huang et al, 2001). Yoneda et al. (2001) has observed that gene expression in human epithelial cells undergo distinct temporal phases when exposed to oxidants such as hydrogen peroxide and smoke. Increased expression of anti-apoptotic genes was observed in epithelial cells one hour after exposure to smoke or hydrogen peroxide (Yoneda et al, 2001).

Similarly, Huang et al (2001) has shown in Sprague-Dawley rats exposed to cisplatin for 7 days, specific gene expression changes associated with temporal regulation of the damage/repair process. The genes altered by cisplatin were subdivided into eight functional categories that included calcium metabolism, oxidative stress, cell cycle regulators, transporters, and metabolism. Temporal regulation was observed as a gradient of responses involving different gene expression patterns associated with the different functional gene categories (Huang et al, 2001). These studies indicate that the genetic cellular response to toxicant exposure can be categorized into distinct phases in which cells first respond to the initial exposure by attempting to counteract the chemical-induced effects on cellular metabolism,

followed by mechanistic counter-active responses that ensure continued cellular function and survival during and after toxicant exposure.

Furthermore, Lu et al. (2001) used gene expression arrays to study toxic effects of arsenic on human populations in China exposed to high levels of arsenic through arsenic-laden coal used for heating and cooking. Liver biopsies from chronically exposed individuals showed increased gene expression for numerous cell cycle regulators and transcription factors involved in oncogenesis and differentiation.

In a study by Thomas et al. (2002), age-related changes in gene expression patterns in rat livers were assessed using DNA microarrays, and compared to changes in human livers. Concomitant increases in the expression of important antioxidants and detoxifying genes were noted in the livers of both rats and humans (Thomas et al, 2002). The induction pattern suggests a complex link between changing hepatic detoxification/redox capability and senescence. Although this study is not directly related to toxicology, the results of this study demonstrate the applicability of using gene expression arrays to characterize mechanisms of action specific to a cellular process.

Real-Time Quantitative Polymerase Chain Reaction (RT-QPCR)

Due to inherent variations from technical and biological sources in microarray experiments, the gene expression results need to be validated using RT-QPCR or other alternate techniques such as northern or western blot. RT-QPCR method allows for the quantitation of gene expression from small

starting mRNA concentrations. It also used to assess discrepancies between or due to hybridization differences in microarray experiments (Goodsaid et al, 2004). RT- quantitative PCR (Overbergh et al, 2003; Grove, 1999) provides a sensitive, reproducible and accurate method for determining mRNA or total RNA levels in tissues or cells. The method is based on the detection of a fluorescent signal produced and monitored during the amplification process, without the need for post-PCR processing (Overbergh et al, 2003).

Quantification of the transcription levels of these genes plays a central role in the understanding of gene function and of abnormal alterations in regulation that may result in a disease state. As a result, the RT-QPCR technique is now widely applied in quantitation of relative gene expression, single nuclear polymorphism detection, cancer and other disease diagnostics, pathogen detection, etc. It is based on the detection of a fluorescent signal produced and monitored during the amplification process, without the need for post-PCR processing.

Various researchers have used RT-QPCR to validate gene expression data obtained using microarray or other large scale toxicogenomic methods (Baker et al, 2004; Heid et al, 1996; Bustin et al, 2002; Burczynski et al, 2001; Grove et al, 1999). Recently, RT-QPCR has been employed to determine reliability and interlaboratory reproducibility of microarray data (Baker et al, 2004). As part of a series of studies conducted by the International Life Sciences Institute Health and Environmental Science Institute Technical Committee on the application of genomics to mechanism-based risk assessment, the

biological response in rats to various hepatotoxicants was investigated. In one of these studies, Baker et al (2004) examined clofibrate-induced gene expression changes in rat liver using membrane cDNA arrays and RT-QPCR. Quantitative RT-PCR was used to confirm modulations for a number of peroxisome proliferator marker genes (Baker et al, 2004). In spite of minimal variability in the quantitative nature of microarray data due experimental procedures and data analysis, the RT-QPCR validation studies indicate the potential for gene expression profiling to identify toxic hazards by the identification of mechanistically relevant markers of toxicity.

Signal Transduction Mechanisms of Genes Selected for RT-QPCR

The candidate genes for Real-time Reverse Quantitative PCR were selected based on their association with health effects or involvement in disease pathways. A summary of each of these genes are given below.

Glutathione-s-Transferase yc2 (Gst-yc2)

Gst-yc2 belong to a family of enzymes that play an important role in detoxification by catalyzing the conjugation of many hydrophobic and electrophilic compounds with reduced glutathione (Klaassen, 1995). *Gst* increases the rate of glutathione conjugation of xenobiotics through deprotonation of GSH to GS^- by an active site tyrosinate ($Tyr\text{-}O^-$), which functions as a general base catalyst (Klaassen, 1995). *Gst* is a polymorphic gene encoding active, functionally different proteins, thought to function in

xenobiotic metabolism and play a role in susceptibility to cancer and other diseases. It expression and activity is regulated by a wide range of xenobiotics and drug compounds such as aflatoxin-b1, coumarin, oltipraz, trans-stilbene oxide, PAHs, and other compounds (Ingenuity Pathway Analysis, Ingenuity Inc, CA). The enzyme is present in most tissues, with high concentrations in the liver, intestine, kidney, testis, adrenal, and lung, where they are localized in the cytoplasm and endoplasmic reticulum (Klaassen, 1995; Ingenuity Pathway Analysis, 2004). The *Gst-yc2* subunit is a member of the Gst apha (*Gsta*) family of proteins; rats express four *Gsta* subunits (ya1, ya2, yc1, yc2). The rat *Gsta* is the major *Gst* family in the liver and kidney. In rats, the individual members of the alpha and mu class of *Gst* are inducible by 3-methylcholanthrene, anti-oxidants, corticosteroids, and phenobarbital (Klaassen, 1995; Bartosiewicz et al, 2001). Induction is usually associated with increased levels of mRNA due to transcriptional activation of the gene encoding a subunit of *Gst*. The enhancer regions of the genes encoding some of the rat *Gst* families have been shown (Bartosiewicz et al, 2001) to contain a xenobiotic (dioxin)-responsive element (XRE).

Electrophiles in cells/tissues are potentially toxic and can bind to critical nucleophiles, such as proteins and nucleic acids, and cause cellular damage and genetic mutations. Therefore, conjugation with glutathione represents an important detoxication. Over-expression of *Gst* is associated with resistance to some toxic compounds (Ingenuity Pathway Analysis, Ingenuity Inc, CA).

Ingenuity Pathway knowledge (Ingenuity Pathway Analysis, Ingenuity Inc, CA) base analysis shows that the *Gst-yc2* gene regulates the ErK/MAP kinase kinase signaling pathway. The microarray data revealed that this pathway is regulated by BaP exposure. Induction of MAP kinase kinase genes was observed in the microarray data. In addition, the expression of *Gst* is also regulated by the large T antigen, *TGM2, fos, EGF* growth factors and other proteins.

Cytochrome p450 (Cyp1a1 and *Cyp1a2)*

Cytochrome p450 are among the most versatile of the phase 1 biotransforming enzymes in terms of the sheer number of xenobiotics they detoxify or activate to reactive intermediates (Klaassen, 1995). The highest concentration of cytochrome p450 enzymes involved in xenobiotic biotransformation is found in liver endoplasmic reticulum but they are also present in virtually all tissues (Klaassen, 1995). The liver microsomal p450 enzymes play a very important role in determining the intensity and duration of action of drugs, and they also play a key role in the detoxication of xenobiotics (Barouki and Morel, 2001; Klaassen, 1995; Chen and Tukey, 1996). In addition, cytochrome p450 enzymes in liver and extrahepatic tissues play important roles in the activation of xenobiotics to toxic and/ or tumorigenic metabolites (Klaassen, 1995). Furthermore, microsomal and mitochondrial p450 enzymes play key roles in the biosynthesis or catabolism of steroid hormones, bile acids, fat-soluble vitamins, fatty acids and

eicosanoids, which shows the catalytic versatility of cytochrome p450 (Klaassen, 1995). All cytochrome p450 enzymes are heme-containing proteins and the basic reaction they catalyze is monoxygenation in which one atom of oxygen is incorporated into a substrate, and the other is reduced to water with reducing equivalents derived from NADPH (Klaassen, 1995, Hall and Grover, 1990).

Cyp1a1 and *Cyp1a2* are monooxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids (Ingenuity Pathway Analysis, Ingenuity Inc, CA). The enzymes endogenous substrate are unknown, however, they are able to metabolize PAHs to carcinogenic intermediates. They are found in the sub-cellular locations such as the cytoplasm, endoplasmic reticulum, microsome, and mitochondrial membranes.

Species differences exist in the relative expression of *Cyp1a1* and *Cyp1a2* in the livers (Ingenuity Pathway Analysis, Ingenuity Inc, CA). For instance, in rats *Cyp1a1* is preferentially used in catalyzing xenobiotics such as in the O-dealkylation of 7-ethoxyresorufin, whereas *Cyp1a2* preferentially catalyzes the O-dealkylation of 7-methoxyresorufin (Klaasen, 1995).

Ingenuity Pathway knowledgebase analysis of the data indicates that their expression is regulated by omeprazole, BaP, Genistein, LPS, and other compounds. Their activity is regulated by several compounds such as tetrachlorodibenzodioxin (TCDD), Ah receptor, 3-methylycholanthrene, 1-aminobenzotriazole, and others. In addition, several of these compounds also

regulate their transcription, quantity, accumulation, and are inhibited by compounds such as tryptamine.

B-cell Lymphoma 2 (Bcl-2)

Bcl-2 is an integral inner mitochondrial membrane protein that blocks the apoptotic death of some cells such as lymphocytes (Brown et al, 1990; Davila et al, 1995). Constitutive expression of *Bcl-2*, such as its translocation to Ig heavy chain locus, is thought to be a cause of follicular lymphoma (Ingenuity Pathway Analysis, Ingenuity Inc, CA). Two transcripts variants, produced by alternate splicing, differ in their C-terminal ends. Variant 1 is represented by a 5.5 kb mRNA and includes most of exon 1 and exon 2. A 3.5 kb mRNA represents variant 2; it includes all of exon 1 and lacks exon 2 sequence (Ingenuity Pathway Analysis, Ingenuity Inc, CA). It is involved in the regulation of cell cycle (anti-apoptosis), humoral immune response, cell growth and/or maintenance, and negative regulation of cell proliferation (Davila et al, 1995; Salas and Burchiel, 1998). It is found in different sub-cellular locations such as cell surface, cytoplasm, endoplasmic reticulum, intracellular membranes, mitochondrial membrane, nuclear envelope, and other locations (Ingenuity Pathway Analysis, Ingenuity Inc, CA). It expression is regulated by tumor necrosis factor (*TNF*), glutamate, *myc*, growth hormone releasing peptide-6, Interleukin-1beta (*IL-1β*), and other factors (Ingenuity Pathway Analysis, Ingenuity Inc, CA). For instance, the activity of *Bcl-2* gene during programmed cell death is believed to depend on collaboration (Alberts et al, 1994) between

myc and the *Bcl-2* gene. If myc alone is overexpressed, cells are driven round the division cycle in appropriately, but no cancer results because the progeny of such abnormally forced divisions are programmed to die (Alberts et al, 1994). If *Bcl-2* is overexpressed at the same time, however, the excess progeny cells survive and proliferate; *Bcl-2* can act as oncogenes because the *Bcl-2* protein inhibits programmed cell death. Phosphorylation of *Bcl-2* is regulated by *MAPK* and *NGF* beta. Other factors also regulate its transcription and quantity in cells (Alberts et al, 1994).

Tumor Suppressor Protein 53 (Tp53)

p53, a nuclear protein, plays an essential role in the regulation of cell cycle, specifically in the transition from G0 to G1 (Jeffy et al, 2000; Alberts et al, 1994, Binkova et al, 2000; Ingenuity Pathway Analysis, Ingenuity Inc, CA). It is found in low levels in normal cell, however, in transformed cell lines it is expressed in high amounts, and can contribute to transformations and malignancies. Also, it is a DNA-binding protein containing DNA-binding, oligomerization and transcription activation domains (Ingenuity Pathway Analysis, Ingenuity Inc, CA). In addition, it is postulated to bind as a tetramer to a *p53*-binding site and activate expression of downstream genes that inhibit growth and/ or invasion, and thus function as a tumor suppressor. Mutants of *p53* frequently occur in a number of different human cancers; they fail to bind the consensus DNA-binding site, and therefore cause the loss of tumor suppressor activity (Ramet et al, 1995). It sub-cellular locations are in DNA

replication foci, centrosome, chromatin, cytoplasm, nucleus, and others (Ramet et al, 1995). It regulates the expression of *Bax, Jun, NFkB1*, Insulin-like growth factor binding protein (*IGFBP3*), *Creb* and other protein products (Ingenuity Pathway Analysis, Ingenuity Inc, CA). In addition, it is regulated by interleukin-3, *NFkB, ATM*, MAPkinase and other factors. As transcription factor, p53 is involved in various biological processes such as cell cycle checkpoint, DNA repair, DNA recombination, apoptosis, DNA damage response, and negative regulation of cell cycle. It is associated with tumorigenesis, immunological diseases, metabolic diseases, and other genetic disorders (Ingenuity Pathway Analysis, Ingenuity Inc, CA). Also, Ramet et al (1995) has shown that *p53* expression is correlated with BaP-DNA adducts in carcinoma cells. Similarly, BaP has been shown to activate the human p53 gene through induction of nuclear factor kappa B (*NFkB*) activity.

Interleukin-1 beta (IL-1β)

IL-1β is a cytokine produced by activated macrophages as a proprotein, which is proteolytically processed to its active form by caspase 1 (Ingenuity Pathway Analysis, Ingenuity Inc, CA). This cytokine is an important mediator of the inflammatory response, and is involved in a variety of cellular activities, including cell proliferation, differentiation, and apoptosis. *IL-1β* induces cyclooxygenase-2 in the central nervous system and thereby contributes to inflammatory pain hypersensitivity. Together with eight other interleukin 1

family genes, they form a cytokine gene cluster on chromosome 2 (Ingenuity Pathway Analysis, Ingenuity Inc, CA). The sub-cellular locations are in axons, cell surface, and plasma membrane. In terms of molecular processes, the expression of *IL-1β* is regulated by a wide range of other genes such as *IL-6, FOS, NFKB1A, IL1RL1*, malic enzyme, *IGFBP4* and others. Also, it regulates the binding, transcription, activation and degradation of several other genes. Furthermore, it expression is also regulated by compounds such as glucocorticoids, prostaglandins, *TNF*, injury and others. PAH have been shown to cause an increase in intracellular calcium levels (Mounho et al, 1997) and cause immunosuppressive effects through a mechanism that involves altered cytokine production/ or disruption of cell surface membrane signaling (Davila et al, 1995).

CHAPTER III

MATERIALS AND METHODS

Chemicals

Benzo[a]pyrene (>99%pure) was purchased from Sigma Chemical (St Louis, MO) and used to prepare diets. All other chemicals and solvents used in these experiments were reagent grade. Reagents for mRNA extraction, Affymetrix and Real-Time PCR gene expression analysis were as recommended by the manufacturer and are listed in the methods sections for those operations.

Chemical Exposures

Male Fischer-344 rats, 6- weeks post-weaning and weighing 140-180g, were obtained from Charles River Laboratories (Mattawan, MI). This study was conducted under federal guidelines (Principles of Laboratory Animal Care; NIH publication no 85-23, revised 1985 and Guiding Principles in the Use of Animals in Toxicology) for the use and care of laboratory animals and the protocols (99-09-01 and 03-04-02) were approved by the Institutional Animal Care and Use Committee (IACUC) and the Environmental Safety Office of Western Michigan University. The animals were individually housed in hanging, stainless steel, wire mesh metabolic cages under controlled conditions of temperature (20°C $\pm$ 2°C) and light cycle 12 hr. light/12 hr. dark with free access to water and a defined diet throughout the study. Animals

were randomized by weight, assigned a unique identification number and assigned into an experimental diet treatment group.

Preparation of Food

A standard rat pellet food (Rodent Diet 5001, Lab Diet, St. Louis, MO) was obtained from the Western Michigan University Animal Care Facility. A general procedure developed in the lab (Chiang, 2001) was used to prepare BaP-spiked rat food pellets. BaP was dissolved in toluene to produce a stock solution of 200mg/ml. 2000g of pellet feed, 300ml of 95% ethanol and an aliquot of stock solution (amount varied depending on final concentration desired in the food) were mixed in 4000ml polypropylene bottle and incubated at room temperature on a rotating shaker for 4hrs. The food was then dried in a vacuum oven at 80°C for 48 hrs to evaporate the solvent. The same process was used for preparing the control food, except that no BaP was added to the mixture. Tests confirm that these conditions remove all of the solvent carriers and further that the toxicant is uniformly distributed in the food pellets using this method.

Dose Verification

The final concentration of BaP in the food pellets was determined using high Performance Liquid Chromatography (HPLC). For the HPLC assay, a Hewlett Packard 1090 HPLC system equipped with a DAD detector and a Zorbax SB-C8 column (150 x 4.6 mm, 5μm) (Hewlett Packard, CA) was used. The

detector was set to monitor the absorbance of BaP at 300 nm. The flow rate was 1.0 ml/min and the injection volume was 10μL. Formic acid was used a solvent A and 0.2% formic acid in acetonitrile was used as solvent B. A linear gradient was run as follows: 0 min, 60% A, hold for 1min; 7min, 100 % A, hold for 3 min; 10.1 min, 60 % A. BaP eluted at 7.1 minutes. Sample (five replicates) preparation for HPLC was done using 2 to 5 grams of dose feed soaked in a diluent containing 70/20/10 (%v/v/v) acetonitrile: methanol: acetone with constant shaking for 24 hrs and then filtered through a 0.22 μm glass fiber filter. A calibration curve spanning the range of expected concentrations was obtained through analysis of four concentration levels of BaP standards prepared from the same BaP stock in triplicates.

Dietary Exposures

Two independent studies were carried out (Acute and Sub-chronic); A 2-week acute exposure and 12-weeks sub-chronic exposure. For both studies, animals were randomly assigned into control or treatment groups (study 1, n = 3 per group; study 2, n = 6 per group).

In the first study (Acute), animals were exposed to either control diet, 0.01mg/g-diet or 0.1 mg/g-diet for 2 weeks. In the second study (sub-chronic), animals were exposed to either control diet, 0.01mg/g-diet, 0.1mg/g-diet or 1.0 mg/g-diet for 12 weeks.

Animal weights were recorded two-three times a week. The amount of the diet consumed was estimated by measuring the amount of the food retained

in the feeding jar at the end of each feeding period, the amount spilled, and subtracting the values from the amount of food given at the beginning of the feeding period. Over the entire exposure period, the total food intake for each treatment group per week was calculated by summation of the measured food intake from each feeding period in that week. These values were then compared against the food intake data of control rats and represented as percentage gain or loss in dietary intake relative to controls.

Animal Sacrifice and Tissue Isolation

At the end of each experimental time period, rats were euthanized using CO_2. liver, lungs, kidneys, thymus, and spleen were excised, weighed (fresh weight) and flash frozen in liquid nitrogen for further analysis. Tissue weights were normalized to the body weights of each animal to obtain actual differences in tissue weight.

The liver tissue was excised and split into two pieces; one half for microarray analysis and one half for Real-time PCR analysis. Tissue isolated for gene expression was immediately frozen in liquid nitrogen and stored at -80° C until RNA isolation.

Messenger RNA (mRNA) Isolation

Liver tissue was disrupted in liquid nitrogen using a mortar and pestle followed by homogenization in extraction buffer (Ambion Inc, Austin, TX) with a Dunce homogenizer. mRNA was isolated using a Poly A pure kit (Ambion

Inc., Austin, TX) and carried out according to the manufacturer's specifications. The purity and concentration of mRNA was determined by gel electrophoresis and UV spectrometry (Gene Quant Pro, Amersham)

Affymetrix GeneChip Analysis

Samples were prepared for microarray analysis according to Affymetrix GeneChip ® protocols (Affymetrix, Inc., Santa Clara, CA). The Affymetrix rat toxicology array (RG-U34A) used in these studies consist of 8,800 genes and expressed sequence tags (EST) in the rat UniGene database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene).
Experimental procedures including double-stranded complementary DNA (cDNA) synthesis and biotinylated complementary RNA (cRNA) preparation were conducted as recommended in the Affymetrix GeneChip Expression Analysis Technical Manual (Affymetrix, Inc., 2003).

Double-Stranded cDNA Synthesis

For double strand complementary DNA (cDNA) synthesis, an aliquot containing 10-15 μg mRNA was first hybridized with 100 pmol T7-(dT) 24 primer [5'-(biotin)-GTCGTCAAAGATGCTACCGTTCAGCA-3'] (GENSET Corp, La Jolla, CA). Primer hybridization was completed in a 20-µl reaction containing a final concentration of 10mM dithiothreitol (DTT), 500mM each of deoxyribonucleoside triphosphate (DNTP) mix, and 1x first strand cDNA buffer. The reaction was incubated at 42°C for 2 min followed by addition of

400-600 U superscript II reverse transcriptase (Gibco Life Technologies, Rockville, MD) to synthesize first- strand cDNA. After 1 hr, second-strand cDNA synthesis was carried out by adding 200 μM each of dNTP, 10 U *Escherichia coli* (*E. coli*) DNA ligase, 40 U *E. coli* DNA polymerase I, 20 U *E. coli* RNase H, and 1x second-strand reaction buffer in a 1.5 mL volume and incubated at 16° C for 2 hr. T4 DNA polymerase was added at the end of the reaction for an additional 5 min and soaked in 10 µl 0.5 M EDTA (Sigma-Aldrich, MO). Phase-Lock Gel (Eppendorf Scientific, Inc., Westbury, NY) extraction with phenol/chloroform followed by ethanol precipitation was subsequently performed to clean up the double-stranded cDNA. The cDNA pellet was resuspended in 12 μl Diethyl pyrocarbonate (DePC) treated RNase-free water.

Biotin-Labeled cRNA Preparation and In Vitro Transcription

Biotin-labeled cDNAs were generated from the cDNA samples by *in vitro* transcription with T7 RNA polymerase, Bioarray High Yield RNA Transcript labeling kit (Affymetrix, Inc). Briefly, 3.3-5 μl double-stranded cDNA was mixed gently with 4 μl each of 10x high-yield reaction buffer, 10x biotin-labeled ribonucleotides, 10x DTT, 10x RNase inhibitor mix, and 2 μl of 20x T7 RNA polymerase provided by the kit and incubated at 37° C for 4-5 hr, with gentle mixing every 30 min. Labeled cRNA was then cleaned with RNeasy Mini kit (Qiagen, Inc, Valencia, CA) to obtain an accurate quantification of the labeled cRNA. 20 μg of labeled cRNA was then fragmented to 35-200 bp with

8 μl 5x fragmentation buffer containing 200 mM Tris-acetate, pH 8.1; 500 mM potassium acetate and 150 mM magnesium acetate in a total volume of 40 μl for 35 min at 94° C. Before hybridization onto GeneChip Array, the quality of labeling and fragmentation was verified on agarose gel.

Hybridization to GeneChip Array

After labeling and fragmentation, 15 μg fragmented biotinylated cRNA was hybridized to the Affymetrix GeneChip Array in a 300 μl cocktail containing 5 μl of 3 nM control oligonucleotide B2, 15 μl 20x eukaryotic hybridization controls, and 150 μl 2x hybridization buffer provided in the GeneChip Eukaryotic Hybridization control kit (Affymetrix, Inc.) together with 3 μl 10 μg/μl herring sperm DNA and 3 μl 50 μg/μl acetylated bovine serum albumin (BSA). The hybridization was carried out at 45° C and 60 rpm in GeneChip hybridization oven 640 (Affymetrix, Inc.) for 16 hr. Prior to hybridization to Affymetrix RG-U34A arrays, sample was hybridized to Affymetrix test3 Arrays for quality control.

Washing, Staining, and Scanning the Array

After hybridization, the washing and staining procedures were carried out on a GeneChip Fluidics station 400 in conjunction with Affymetrix Microarray Suite 5.0 software (MAS 5.0; Affymetrix, Inc.). Briefly, the array was first washed with 10 cycles of 2 filling and draining cycles in a non-stringent buffer containing 6x SSPE (52.9 g sodium chloride, 8.28 g sodium phosphate,

monobasic, 2.82 g EDTA, pH 7.9) and 0.01 % Tween 20 at 25°C, followed by 4 cycles of 15 filling and draining cycles in a stringent buffer containing 100 mM MES [2-(N- morpholine) ethanesulfonic acid], 0.1 M Na^+ and 0.01% Tween 20 at 50° C. The probe array then was first stained with a 600 μl streptavidin-phycoerythrin (SAPE) solution containing 1x MES stain buffer (100 mM MES, 1 M Na^+, and 0.05 % Tween 20), 2 μg/μl acetylated BSA, and 10 μg/μl SAPE (Molecular Probes, Inc., Eugene OR) for 10 min at 25° C. After the first staining the array was washed with 10 cycles of 4 filling and draining cycles in non-stringent buffer at 25° C. The stained signals were then amplified in a 600 μl antibody solution containing 1x MES stain buffer, 2 μg/μl acetylated BSA, 0.1 μg/μl normal goat IgG, and 3 μg/ μl antistreptavidin biotinylated antibody (Vector laboratories, Inc) for 10 min at 25° C, followed by a second SAPE staining at the same temperature for another 10 min. Finally, the probe array is washed for 15 cycles of 4 filling and draining cycles in a non-stringent buffer and scanned by the Agilent GeneArray Scanner (Affymetrix, Inc.) at an excitation wavelength of 570 nm. After scanning, each image was inspected for major chip defects or abnormalities in hybridization signal as a quality control and analyzed using MAS 5.0.

The scanner generates an image of the array by exciting each feature with its laser, detecting the resulting photon emissions from the fluorescently labeled RNA that has hybridized to the probes in the feature, and converting the detected photon emissions into a 16-bit intensity value. The images generated by the scanner are then ready for analysis.

For the GeneChip analysis, two rats (n=2) were assayed individually from each treatment group for both 2- and 12-weeks time points. The choice of two animals per group for GeneChip Analysis is based on, 1) gene expression validation studies that indicate one chip per treatment group may be sufficient for expression analysis, 2) n of 2 is considered cost-effective in terms of data analysis since p (number of genes) is large. Each sample (treatment or control) for analysis was first run on an Affymetrix test3 array before hybridization to Affymetrix RG-U34A array. Test chip contains probes corresponding to commonly expressed genes from human, rat, mouse and yeast genomes. Probes derived from 5', middle and 3'-portions of genes and can identify insufficient or degraded RNA target.

In order to assess the reproducibility of the GeneChip Microarray system, two separate samples (control and 0.1mg/g-diet) from the same animal (12-weeks exposed) were run on duplicate chips. This analysis revealed a 96% similarity between the technical replicates from both samples. Therefore, subsequent samples were analyzed using one chip per animal for each treatment group.

Data Analysis

The Affymetrix RG-U34A GeneChip contains 8800 genes and expressed Sequence tags (ESTs) derived from the UniGene database. Each gene is represented by 20 pairs of 20-25 mer sequence probes that contain a perfectly matched probe (PM) and a mismatched (MM) sequence probe that is altered by one base pair. The MM probes serve as controls for background

and cross hybridization signals. Gene Array Laser Scanner software quantifies the expression of each gene by subtracting the signal intensity of the MM from the PM probes.

The Affymetrix Microarray Suite 5.0 (MAS 5.0) software supplied by Affymetrix to process array images from the scanner performs all of the fundamental operations necessary to analyze an array, including (I) image segmentation, (2) background correction, (3) scaling/normalizing arrays for array-to-array comparisons, (4) calculation of statistics to indicate whether a gene transcript is present or absent and (5) calculation of statistics to indicate whether a gene transcript is differentially expressed. The software uses a variety of parameters to determine the expression level of each RNA molecule. Absolute call determines whether RNA is present or absent, and difference call determines the change between baseline and experimental samples. The software also computes a signal log ratio that quantifies the change in signal intensity between two samples (e.g. control versus treatment or control versus control). These files are designated CHP files.

In this project, two GeneChips were processed for each dose and time point. For analysis, comparative CHP files were produced that used the control sample (at each time point) as a baseline and treatment groups as the experimental sample. Each experimental dose (n=2) was compared with each time-appropriate control (n=2), resulting in a total of 4 pair wise comparisons for each dose and exposure time interval. In order to minimize false positives, only those transcripts that had an increase or decrease of ≥3-

fold in at least 3 out of the 4 pair wise comparisons were selected for interpretation. Additional data analysis to generate exploratory statistics, scatter plots, filter genes based on various parameters (fold changes, signal, signal-log ratios, count/percentage, Average Difference Intensity), clusters (*Hierarchical and Self-organizing maps*), and to determine genes that were unique or common between experiments was done using Affymetrix Data Mining Tool (DMT 3.0) and Spotfire decision site for functional genomics (Spotfire, Inc. CA). Smart Draw program (Smart Draw.com, CA) was used to plot Venn diagrams that showed genes that are in common and unique between treatment groups.

Cluster Analysis and Biological Processes Identification

Extracting biological process information from microarray data has been of recent interest and some of the most recent analysis includes k-means clustering, hierarchical clustering, self-organizing maps (SOM), and Principal component analysis (PCA) (Eisen et al, 1998; Waring et al, 2001; Tamayo et al, 1999; Yeung and Ruzzo, 2001). Because of the large number of genes and complexity of biological networks, clustering is a useful exploratory technique for analysis of gene expression data. Each of these techniques has their limitations and therefore, a combination of two or more of these techniques are suggested to represent a given data set. For hierarchical clustering of the data, we used the Spotfire decision site program for functional genomics to cluster the data set into genes that perform similar

function or were induced by BaP in a similar pattern. The hierarchical clustering measures similarity between genes based on expression profiles across conditions. This measurement requires a definition of similarity that is provided by a metric. The most common metrics are Euclidean distance, which measures how different two profiles are taking into account the magnitude of signals. Pearson correlation metric, on the other hand, measures the shape or direction and association between two different profiles. We used the weighted average linkage method for hierarchical clustering of the genes and Euclidean distance as the similarity metric. The average signal intensity of each gene in each treatment group (8800) was normalized to the control signal. The resulting normalized average signal ratios were log transformed using log base 2. The log transformed signals were used as the input values for hierarchical clustering of the data in Spotfire decision site for Functional Genomics (Spotfire, Inc, CA). A threefold log ratio was used to color the resulting hierarchical cluster. Genes with a positive shift or increase equal or greater than threefold were shown with increasing intensity of red. Genes showing a negative shift or decrease equal to or less than negative log of threefold are colored green. The relationships among genes are represented by a dendrogram whose branch length reflects the degree of similarity between the genes. From the resulting dendrogram of the hierarchical cluster, 10 small clusters were created to show the behavior of genes within those clusters. In order to visualize the orientation of these genes in N-dimensional space and get a clear picture of the behavior of the

genes in each cluster, principal component analysis was done on the 10 clusters extracted from the hierarchical cluster.

Self-organizing maps (SOM) have recently been proposed as an alternate unbiased method for extracting patterns; however, there are subtle issues such as choosing the initial geometry of nodes and related parameters. Therefore, a combination of both hierarchical clustering and SOM were utilized to interpret the results. SOM clustering was performed using the Affymetrix Data Mining Tool software 3.0. For SOM analysis, probe sets were considered across experiments (n=2 for treatments) at each time-point. Average signal intensity of each probe across experiments was used to compare relative gene expression across experiments. The following parameters were used; threshold signal was set at 20-20,000; row variation filter: max/min = 3 and max – min = 100; grid geometry: rows (3) x columns (2); initialization: random vectors; Neighborhood: bubble; neighborhood width: 5-0.02; learning rate (α): 0.1-0.05. The details and description of these parameters can be found in Affymetrix data mining tool user's guide, version 3.0. The row expression filter eliminated genes that did not change significantly across the experiments and the genes that passed through the filter were fed to the SOM algorithm.

For Principal component analysis (PCA), we utilized the Spotfire Decision site for functional Genomics (Spotfire Inc, CA) to analyze the data. PCA is a classical technique to reduce the dimensionality of the data set by transforming to a new set of variables (Principal components) to summarize

the features of the data. These summary variables can be used for visualization or for more complex statistical modeling. The first aspect in PCA (Yeung and Ruzzo, 2001) is creation of the components, which are weighted averages of the original variables and they are constructed to be uncorrelated with each other and to capture much of the original variability as possible. The second aspect is the selection of the most representative (or principal) components (PC) based on the fraction of variability that is retained using them only. As in cluster analysis, PCA are also based on similarity (distance) matrix between genes or objects. The Pearson correlation matrix commonly used in gene expression analysis was similarity metric employed. Here the first PC is the direction along which there is greatest variation in the data. For example, we have 8800 genes in our array and therefore, the first PC is described by a vector of "loadings" of length J, consisting of weights to be applied to each gene's expression. A loading close to 0 implies that a gene does not vary much across arrays. Large negative and large positive loadings imply that a gene shows considerable variation across arrays in the dataset. The second principal component is created so that it is orthogonal to the first PC. It represents the weighted average across genes that explain the largest amount of variation in the data after controlling for the variation explained by the first PC. As described for hierarchical clustering, Affymetrix data from both time points and treatments was first transformed before importing data into Spotfire. First, the average signal intensity of each gene from the various treatment groups was normalized to the appropriate control signals to obtain

average signal ratios as previously described. Average signals were log normalized to obtain log signal ratios. The algorithms for PCA and many other clustering algorithms work best with log normalized values. The log normalized values for time point 1 (2-weeks) and Time-point 2 (12-weeks) were combined into one table using time as an identifier to separate them. These data was imported into Spotfire Decision site for Functional Genomics and analyzed by PCA.

Biological Processes and Molecular Function Analysis

After clustering and PCA assessments, it is necessary to interpret microarray data to some useful biological endpoints. Biological process and molecular function information for genes was obtained using Affymetrix NetAffix and Gene Ontology program of National Center for Biotechnology Information (NCBI). Other biological process information was obtained from various databases such as KEGG, SWISSPROT, NCBI UniGene and GenBank databases. Ingenuity Pathways Analysis (Ingenuity Systems, CA) and GenMaPP software (University of California Gladstone Institute, CA) for biological pathways identification was used for mapping biochemical pathways associated with various genes. Ingenuity pathways analysis allows researchers to input into its pathways analysis application a set of genes or proteins, the software's knowledge base then dynamically computes a set of relevant pathways. Similarly, GenMaPP is a free computer application designed to visualize gene expression and other genomic data on maps

representing biological pathways and groupings of genes. In addition, Paracel pathworks (Paracel, Inc, CA) was used to easily draw the biological pathways obtained from Ingenuity and GenMaPP.

Real-Time Quantitative PCR (RT-QPCR)

We use the ABI Prism 7700 SDS in combination with Taqman chemistry, using specific sets of primers and an internal fluorogenic probe for each target of interest. In designing the primer sets, special care was taken to avoid co-amplification of contaminating genomic DNA.

Primer and Probe Design

mRNA sequences of 6 rat genes of interest (*cytochrome p450 1a1, Cyp1a1; cytochrome p450 1a2, CyP1a2; Glutathione-s-transferase yc2, Gst-yc2; B-cell lymphoma, Bcl-2; Tumor suppressor protein 53, p53; Interleukin-1beta, IL-1β)* and yeast actin control (internal standard) were downloaded from NCBI GenBank. Real-time quantitative PCR (Taqman) primers and probes for the yeast actin control and rat genes of interest were designed employing Primer Express software (Applied Biosystems). The primary object of this program is to design sets of primers and internal probes that can be run under universal thermal cycling conditions (15 s at 94° C and 1 min at 60° C). The default parameters of the software are therefore set to be very narrow. The required parameters for well-designed primers and probe have been well documented and are built into the program. These parameters include a T_m for the probe

that is 10° C higher than the primers, primer T_ms between 58° C and 60° C, amplicon size between 50 and 150 bases, absence of 5' Gs, and primer length. More specifically for our experiment, we designed primers with no more than 2 Cs or Gs in last five nucleotides of the 3' end; probes with high Cs in sequence and no Gs at 5' of probe. The 5'- and 3'-end nucleotides of each probe were labeled with a reporter flourochrome (6-carboxyfluorescin or FAM) and a quencher dye (6-carboxy-tetramethyl-rhodamine or TAMRA), respectively. Where possible, to avoid co-amplification of contaminating genomic DNA, primers are designed on different exons or intron-exon boundaries.

Taqman Chemistry and Principle

The principle of the Taqman real-time detection is based on the fluorogenic 5' nuclease assay. The Taqman chemistry, which is the key to the detection system based on two principles, 1) the intrinsic 5'-3' exonuclease activity of *Taq* polymerase, and 2) the presence of a dual-labeled Taqman probe. The Taqman probe is designed to anneal to the target sequence between the traditional forward and reverse primers. Also, the probe is created to emit a fluorescent signal only upon cleavage, based on fluorescence resonance energy transfer (FRET). The probe is labeled at the 5' end with a reporter flourochrome (6-FAM) and a quencher flourochrome (TAMRA) added at the 3'end. The probe is designed to have a higher T_m than the primers, and during the extension phase, the probe must be 100 % hybridized for success

of the assay. As long as both fluorochromes are on probe, the quencher molecule stops all fluorescence by the reporter. However, as *Taq* polymerase extends the primer, the intrinsic 5' to 3' nuclease activity of *Taq* degrades the probe, releasing the reporter fluorochrome. The amount of fluorescence released during the amplification cycle is proportional to the amount of product generated in each cycle.

The ABI PRISM 7700 sequence detection system used in this experiment consists of a 96-well thermal cycler connected to a laser and charge-coupled device (CCD) optics system. An optical fiber inserted through a lens is positioned over each well, and laser light is directed through the fiber to excite the fluorochrome in the PCR solution. Emissions are sent through the fiber to the CCD camera, where they are analyzed by the software's algorithms. The software package supplied with the instrument measures the increase in fluorescence emissions in real-time, during the course of the reaction. The increase in fluorescence is directly related to the increase in target amplification. The software calculates the difference between the fluorescence emissions of the product at each time point and fluorescence emissions of the baseline. The cycle number at which the fluorescence emissions of the product rise above the background or baseline is the threshold cycle (C_t). The C_t is determined at the most exponential phase of the reaction (usually 10 times the standard deviation of the baseline) and is more reliable than end-point measurements of accumulated PCR products

used by traditional PCR methods. The starting copy numbers of unknown samples is determined through standard curve plots of C_t values.

Table 1

Sequence of Real-Time PCR Primers and Probe

Gene	Forward Primer	Reverse Primer	Probe
Cyp1a1	ccccggctttctgacaga	atgtcggaaggtctccagga	ctcagctgccctatctggaggccttc
Cyp1a2	tgatgagaagcagtggaagacc	ggccgtgttgtcattggtaag	tttgtgttccgcccagagcggtt
Gst-yc2	gccgatctggagttgatggt	atcttggcaagactcgcctc	ctctattacccctacatgccccctggg
P53	catgagcgttgctctgatg	cagatactcagcatacggatttcc	acggcctggctcctcccaac
Bcl-2	aggattgtggccttctttgagtt	gccggttcaggtactcagtcat	ccctggtggacaacatcgctct
IL-1β	cacctctcaagcagagcacag	gggttccatggtgaagtcaac	tgtcccgaccattgctgtttcctagg
Yeast	tggattccggtgatggtgtt	tcaaaatggcgtgaggtagaga	ctcacgtcgttccaatttacgctggttt

RT-QPCR Reaction and Amplification

All reagents for one-step RT-PCR reactions were purchased from Applied Biosystems, Inc. PCR amplifications are performed on the ABI PRISM 7700 SDS using 96-well microtiter plates. They are performed on a total volume of 25 μl containing 4 ng of mRNA, 300 nM primers (forward and reverse), 200 nM Taqman probes, AmpliTaq Gold DNA polymerase mix (2X) containing AmpliTaq Gold DNA polymerase, dNTPs with dUTP, Passive Reference, Taqman optimized buffer components, and 40X RT enzyme mix containing MultiScribe™ Reverse Transcriptase and RNase inhibitor. All RT-PCR plates

were covered with optical adhesive covers following complete loading of reagents. For each gene of interest and yeast internal control, different primer, probe and poly A RNA concentrations were tested to optimize the PCR amplification. Also, no enzyme (reverse transcriptase) controls were run for each reaction. RT and PCR amplifications were performed in duplicate wells (except for the No enzyme controls), using the following temperature cycles: 30 min at 48° C for the RT hold, 10 min at 95° C (AmpliTaq Gold DNA polymerase activation), and then run for 40 cycles of PCR at 95° C for 15 s (denature), and 60° C for 1 min for annealing and extension.

Normalization

In order to control for variations due to mRNA isolation or RT-PCR conditions, 50 ng of yeast mRNA was spiked into each sample during mRNA extraction. Also, RT-PCR for each gene of interest was run in conjunction with yeast actin for each sample to normalize the expression levels.

Quantification and Data Analysis

To quantify the results obtained by real-time RT-PCR we used the standard curve method. Standard curves for both the yeast internal control and the target genes was generated from serial dilutions of untreated control samples, and used to determine the amount of mRNA for each gene and yeast in all experimental samples. The relative expression level of each gene was computed with respect to yeast mRNA for each treatment group to account

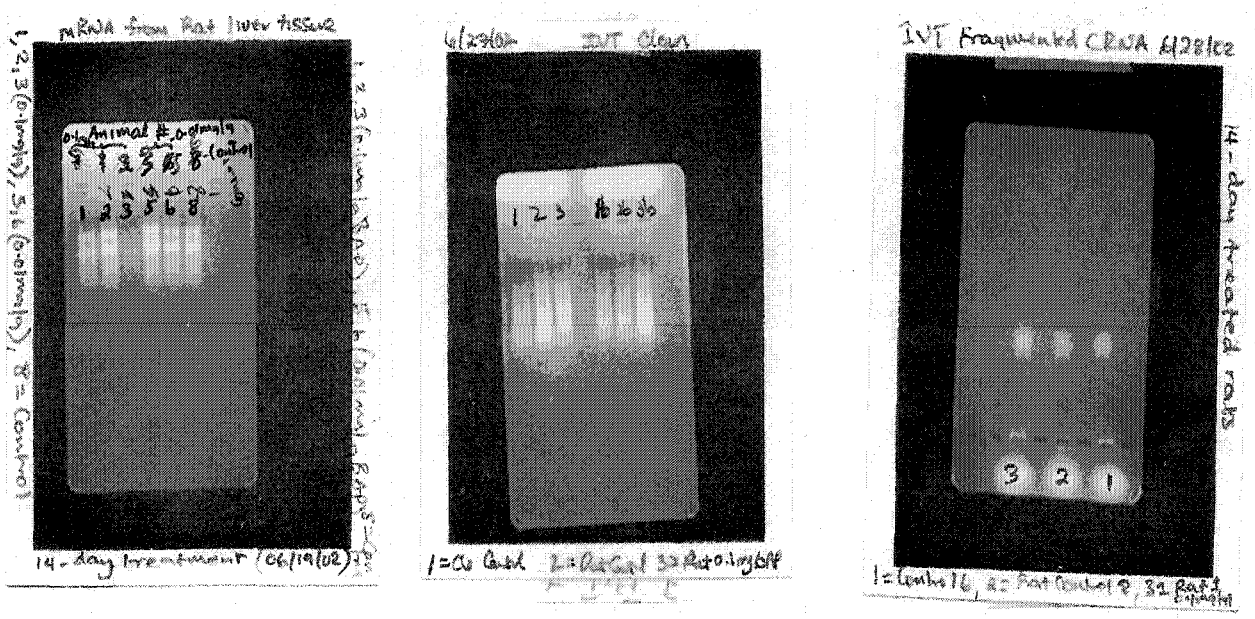
for variation in quality of mRNA or RT-PCR conditions. Gene Expression levels for the treatment groups were compared with control groups using one-way ANOVA F- statistics followed by Bonferroni pairwise post-hoc test. Levene statistic was used to test for homogeneity of variances.

CHAPTER IV

RESULTS AND DISCUSSION

Affymetrix Microarray Analysis

To determine the effects of BaP on the expression patterns of genes in the rat liver, RNA was extracted from rat livers and the results showed that all of the RNA was intact. Representative gels are shown in Figure 5 below.



(a) (b) (c)

Figure 5. Representational Gels of mRNA, cRNA and Fragmented cRNA from Rat Liver Tissue. (a) mRNA from rat liver tissue of 2-weeks exposed animals, (b) cRNA produced from *In Vitro* Transcription of cDNA, and (c) fragmented cRNA.

Gene expression profiles were determined using Rat toxicology RG-U34A microarrays from Affymetrix. Expression changes were reflected in the

differences in signal intensity from array hybridizations. Analysis of the microarray differences in the hybridization signals was accomplished by the use of Affymetrix Microarray Analysis Suite 5.0 software (MAS 5.0). A threefold or greater change in signal intensity was considered a significant change in expression (Figure 6-10). Analysis of the gene expression patterns of rat livers in response to BaP demonstrated a threefold or greater change in the expression of a large number of genes at both 2-and 12-weeks time and exposure levels (Figures 6-10). Figure 6-10 shows scatter plots of differential signal intensity between control and treatment groups (Pairwise comparison) for 2- and 12-weeks exposure. Fold change of 2, 3, and 10 are represented by black lines.

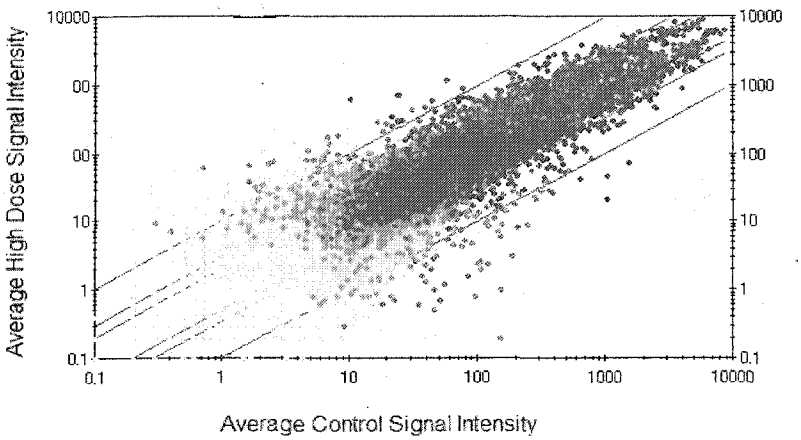


Figure 6. Scatter Plot of Differential Signal Intensity between Control and High Dose (0.1 mg/g of diet) After 2-weeks of Exposure to BaP

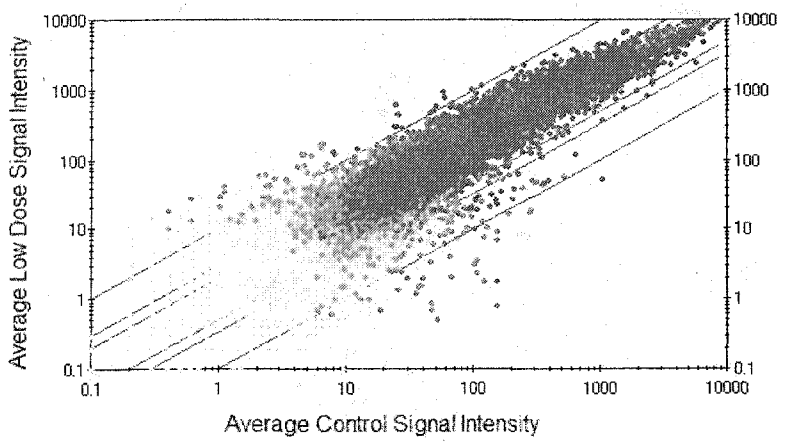


Figure 7. Scatter Plot of Differential Signal Intensity between Control and Low Dose (0.01 mg/g of diet) After 2-weeks of Exposure to BaP

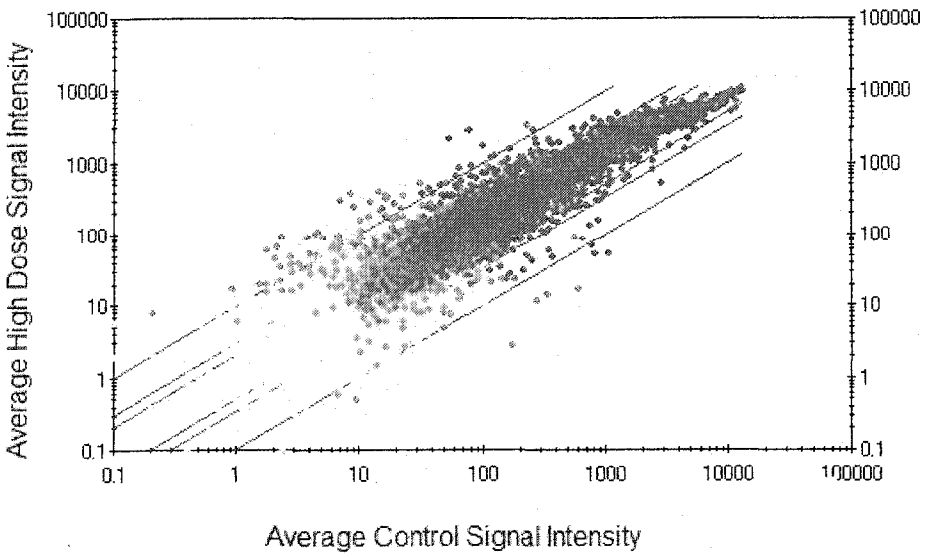


Figure 8. Scatter Plot of Differential Signal Intensity between Control and High Dose (1.0 mg/g of diet) After 12-weeks of Exposure to BaP

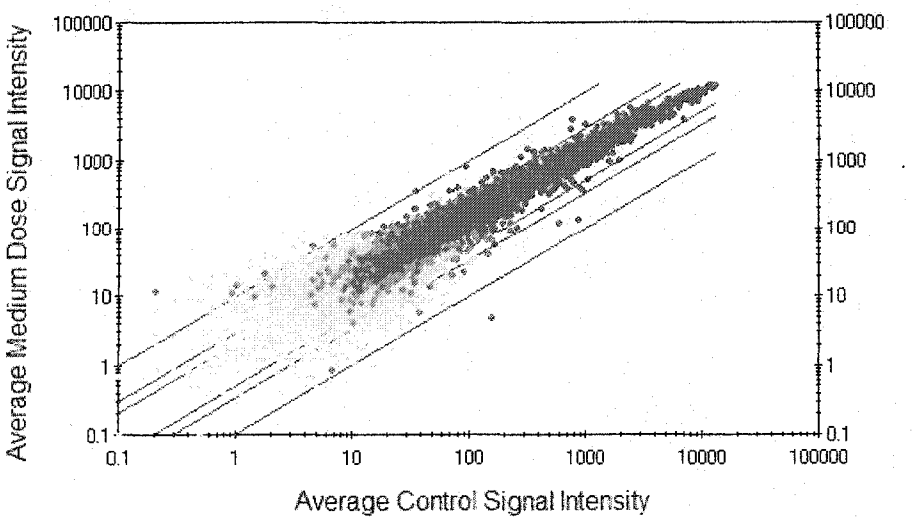


Figure 9. Scatter Plot of Differential Signal Intensity between Control and Medium Dose (0.1 mg/g of diet) After 12-weeks of Exposure to BaP

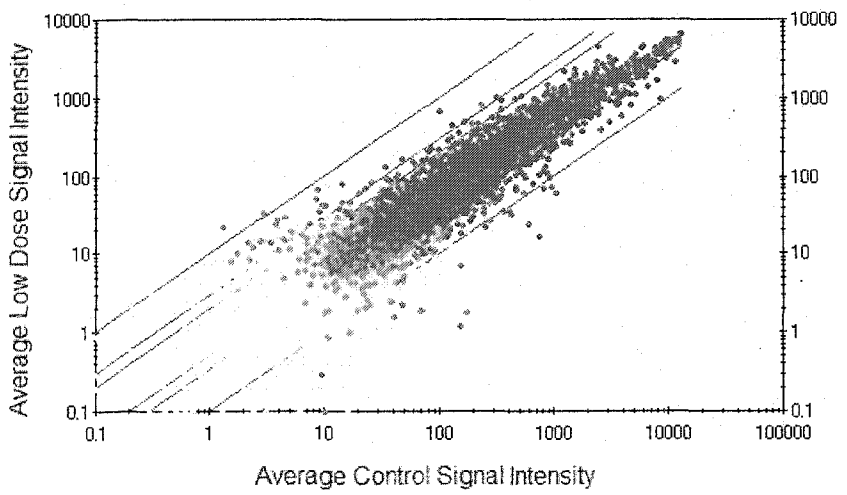


Figure 10. Scatter Plot of Differential Signal Intensity between Control and Low Dose (0.01 mg/g of diet) After 12-weeks of Exposure to BaP

As shown in Figures 6 and 7, after 2-weeks of exposure to BaP, a large number of genes were differentially expressed (up to 10 fold) at both low and high dose exposure. This is in line with several studies characterizing the temporal regulation of gene expression in cells or tissues exposed to various toxicants (Huang et al, 2001; Yoneda et al, 2001). The induction of large number of genes at early time points is explained by the genetic cellular response to toxicant exposure by various cells. The study by Yoneda et al (2001) show that cells first respond to an initial exposure by attempting to counteract the chemical-induced effects on cellular metabolism, followed by mechanistic counter-active responses that ensure continued cellular function and survival during and after toxicant exposure. On the other hand, after 12-weeks of exposure, differential expression of genes was greater for the high and low dose groups compared to the medium dose groups. Some of the variations may also arise from intrinsic sources of biological variability due to inherent genetic variation among individuals within a population and the environmental conditions.

Quality Control and Validation of Affymetrix GeneChip Method

Initial samples were run on duplicate chips in order to determine the chip to chip variability of the Affymetrix gene chips. For this analysis, we used one control sample and one 0.1mg/g-diet BaP treated sample from the 12-weeks exposed animals. For both samples, the chip replicates showed a 96.0% similarity (Table 2) in gene expression. Therefore subsequent samples were

run on single chips. Analysis was carried out to determine the variability between animals from the same treatment groups. The results of this analysis are shown in Table 3. This analysis showed that there was an 80-88% similarity between animals (n=2) from the same treatment group. Table 3 shows that for each treatment group and time point, the animal to animal variability was minimal (≤ 20%).

Table 2

Gene Expression Changes of Technical Replicates

Dose (mg/g-diet)	Exposure Time (Wks)	N	% Similarity
0.1	12	2	96.0
Control	12	2	95.6

This is consistent with other reports (Bartosiewicz et al, 2001) on the validation of the use of microarrays in an *in vivo* model. Bartosiewicz et al. (2001) used various doses of α-napthoflavone, a known inducer of *Cyp1a1* and *Cyp1a2* genes in murine liver, to assess the issues of sensitivity and dynamic range of microarray analysis (spotted arrays) in comparison to Northern blots as well as to investigate sources of variability in the assay (slide to slide, spot to spot, and animal to animal). The results from their work shows that the dynamic range and sensitivity of DNA microarrays is comparable to Northern blot analysis and that variability in the data is largely

due to inter-animal differences rather than analysis-related variations in hybridization or technique.

Table 3

Gene Expression Variability between Animals in Individual Treatment Groups

Dose (mg/g-diet)	Exposure Time (wks)	N	% Similarity
0.01	2	2	85.0
0.1	2	2	82.0
Control	2	2	82.0
0.01	12	2	88.0
0.1	12	2	80.0
1.0	12	2	85.0
Control	12	2	83.0

Exposure to BaP Altered Gene Expression in the Rat Liver

Table 4 summarizes the number of transcripts that were regulated in response to BaP. Comparative CHP files produced in MAS 5.0 that used the control sample (at each time point) as a baseline and treatment groups as the experimental sample were generated. Each experimental dose (n=2) was compared with each time-appropriate control (n=2), resulting in a total of 4 pair wise comparisons for each dose and exposure time interval. In order to minimize false positives, only those transcripts that had an increase or

decrease of ≥3-fold in at least 3 out of the 4 (75-100%) pair wise comparisons were selected for interpretation. After two-weeks, exposure to 0.01mg/g-diet increased expression of 266 genes and decreased expression of 67 genes relative to controls, and two weeks of exposure to 0.1mg/g-diet increased expression of 53 genes and decreased expression of 79 genes. After 12-weeks, a dose-dependent trend occurred in several genes that were up regulated. For example, 28 genes were increased in rats exposed to 0.01mg/g-diet compared to controls, 64 genes increased in rats exposed to 0.1mg/g-diet and 102 genes increased in rats exposed to 1.0mg/g-diet. Furthermore, the expression of a large number of genes were decreased after exposure to BaP for 12-weeks (185 genes decreased in rats exposed to 0.01mg/g-diet; 30 in rats exposed to 0.1mg/g-diet and 76 in rats exposed to 1.0mg/g-diet).

In a similar study with a limited set of genes on a custom array, Bartosiewicz et al (2001) observed induction and repression of genes in liver and kidney in response to intraperitoneal administration of 80 and 100 mg/kg of BaP. As observed in this study, both time and dose were found to have significant effects on the pattern of gene expression of the chemicals tested. Both time and dose are important factors in determining not only which genes are upregulated but also to what extent their expression increased. In this study, induction of specific genes, on the one hand, followed the classic dose-response relationship. On the other hand, the induction of a large number of genes was found to be dose-independent. The dose-response relationships

are not always monotonic (Neumann and Galvez, 2002). Exposures to low, sub-lethal concentrations of toxicants may actually enhance some physiological responses, while exposure to a more elevated concentration of the same toxicant may be inhibitory in a dose-dependent fashion.

Table 4

Total Number of Transcript Altered by Exposure to BaP

Dose	Time	3-Fold Change		2-Fold Change	
Mg/g	Weeks	Increased	Decreased	Increased	Decreased
0.01	2	266	67	1279	168
0.1	2	53	79	181	226
0.01	12	28	185	90	571
0.1	12	64	30	329	100
1	12	102	76	343	223

Dose Comparisons for Genes in Common/Unique for Individual Exposure Time Intervals

Analysis of the genes whose expression was increased or decreased across dose for individual exposure time intervals showed several genes in common and others that were uniquely expressed. Affymetrix DMT 3.0 software tool was used to create and combine probe lists of genes/transcripts that increased or decreased across dose and time points. The Boolean algorithm

'and' combines two different probe lists and produces a probe list of genes that are in common for the two probe lists.

Figures 11 shows Venn diagram plots of genes that are in common or unique across the different doses for the 2-weeks and 12-weeks time points. For the genes whose expression was increased by 0.01mg/g-diet and 0.1mg/g-diet of BaP after two weeks (Figure 11a), 16 genes were in common between these two dose groups. The gene IDs, gene names and molecular function description of these genes is shown in Table 5. These include immune response genes (interleukin-1 beta, *CD24* and *CD74* antigen), anti-oxidants (glutathione-s-transferase yC2), transcription factors (tumor suppressor protein *p53*, pyruvate dehydrogenase, *MIPP65*), transporters, and other genes. Similarly, for the genes whose expression was decreased by 0.01mg/g-diet and 0.1mg/g-diet of BaP after two weeks (Figure 11b), 15 genes were found to be in common. As shown in Table 6, these genes include sodium ion transporters, transcription and DNA binding factors (forkhead box, meprin 1 alpha, platelet derived growth factor receptor), and other notable genes.

On the other hand, for genes whose expression was increased after twelve weeks exposure to 0.01mg/g-diet, 0.1mg/g-diet and 1mg/g-diet of BaP (Figure 11c), 8 genes were in common for all 3 doses. As shown in Table 7, these genes include, the potassium voltage gated ion channels, cell cycle regulators (B-cell lymphoma 2, pericentriolar material 1), Interleukin-1 beta, and glutathione-s-transferase yc2. In addition, between doses (Figure 11c), 5

genes were in common between 0.01mg/g-diet and 0.1mg/g-diet, 4 in common between 0.01mg/g-diet and 1mg/g-diet, and 11 genes in common between 0.1mg/g-diet and 1mg/g-diet.

Table 5

Genes That are in Common to All Doses and Showed Increased Expression (≥ 3.0 fold) After 2-weeks of BaP Exposure

Probe Set ID	Gene Title	Molecular Function Description
AB000098_at	MIPP65 protein	
AF062741_at	pyruvate dehydrogenase phosphatase isoenzyme 2	magnesium ion binding, catalytic activity protein serine/threonine phosphatase activity
AF081366_s_at	potassium inwardly-rectifying channel, subfamily J, member	ion channel activity, voltage-gated ion channel activity, ATP binding
L01507_at	POU domain, class 1	DNA binding, transcription factor activity
L79910_s_at	Sarcosine dehydrogenase	
M21842_at	apolipoprotein B	cholesterol metabolism
E01884_at	Interleukin-1 beta	cytokine activity, interleukin-1 receptor binding
rc_AI058393_s_at	Arg/Abl-interacting protein	intracellular signaling cascade
rc_AI101743_s_at	hydroxysteroid (17-beta) dehydrogenase 4	estradiol 17-beta-dehydrogenase activity, sterol carrier activity, oxidoreductase activity
S82820_at	Glutathione-s-transferase yc2	glutathione transferase activity, transferase
U49062_at	CD24 antigen	defense response
X13058_at	tumor protein p53	DNA binding, transcription factor activity
X14254cds_g_at	CD74 antigen	chaperone activity, immune response
X63253cds_s_at	solute carrier family , member 4	neurotransmitter:sodium symporter activity
Z35654_at	MCF.2 cell line derived	Guanyl-nucleotide exchange factor activity

Table 6

Genes That Are in Common to All Doses and Showed Decreased Expression (≥ 3.0 fold) After 2-weeks of BaP Exposure

Probe Set ID	Gene Name	Molecular Function Description
D90102_s_at		
M22253_at	sodium channel, voltage-gated, type 1, alpha polypeptide	ion channel activity, voltage-gated ion channel activity, voltage-gated sodium channel activity cation channel activity, calcium ion binding
M22254_at	sodium channel, voltage-gated, type 2, alpha 1	Voltage-gated sodium channel activity
M87634_at	forkhead box O1	DNA binding, transcription factor activity
rc_AA875390_at	thioredoxin-like (32kD)	
rc_AI639248_r_at	proline-rich protein	
S43408_g_at	meprin 1 alpha	endopeptidase activity, meprin A activity metallopeptidase activity, zinc ion binding
S56508_s_at	acyl-CoA synthetase long-chain family member 6	magnesium ion binding, catalytic activity long-chain-fatty-acid-CoA ligase activity
S73009_at	synuclein, alpha	
S75275_at	Similar to DEAD-box protein 4 (VASA homolog)	
U10303_at	endothelial differentiation sphingolipid G-protein-coupled receptor 1	rhodopsin-like receptor activity lysosphingolipid and lysophosphatidic acid receptor activity, G-protein coupled receptor a
Z14118cds_at	platelet derived growth factor receptor, alpha polypeptide	protein serine/threonine kinase activity, transmembrane receptor protein tyrosine kinase receptor activity, ATP binding, dimerization

For genes whose expression was decreased after 12-weeks, exposure to 0.01mg/g-diet, 0.1mg/g-diet and 1mg/g-diet of BaP (Figure 11d) showed 3 genes in common for all 3 doses. The three genes in common shown in Table

8 are DNA binding protein N5, glutamate receptor, and the EST rc_AA859372.

Table 7

Genes That Are in Common to All Doses and Showed Increased Expression (≥ 3.0 fold) After 12-weeks of BaP Exposure

Probe Set ID	Gene Name	Molecular Function Description
D89730_g_at	Similar to EGF-containing fibulin-like extracellular matrix protein 1 precursor (Fibulin-3) (FIBL-3) (T16 protein) (LOC305604), Mrna	
J04731_at	potassium voltage-gated channel, shaker-related subfamily, member 2	ion channel activity, voltage-gated potassium channel, cation channel and potassium channel activity, protein binding
M63991_at	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7	serine-type endopeptidase inhibitor activity thyroid hormone transporter activity
S82820_at	Glutathione-s-transferase yC2	Glutathione transferase activty
E01884_at	Interleukin -1 beta	cytokine activity, interleukin-1 receptor binding, growth factor activity
L14680_at	B-cell lymphoma 2	protein binding
rc_AI639037_		
U95920_at	pericentriolar material 1	regulation of cell cycle

Between doses (Figure 11d), 23 genes were in common for 0.01mg/g-diet and 1mg/g-diet, 5 were in common between 0.01mg/g-diet and 0.1mg/g-diet, and 5 genes in common between 0.1mg/g-diet and 1mg/g-diet.

Table 8

Genes That Are in Common to All Doses and Showed Decreased Expression (≥ 3.0 fold) After 12-weeks of BaP Exposure

Probe Set ID	Gene Name	Molecular Function Description
L31882_at	DNA binding protein (N5)	double-stranded DNA binding transcription factor activity
M36419_s_at	glutamate receptor, ionotropic, 2	receptor activity, ionotropic glutamate receptor activity, alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate selective glutamate receptor activity, transporter activity, ion channel activity glutamate-gated ion channel activity potassium channel activity
rc_AA859372		

Exposure Time Interval Comparisons for Genes in Common and Unique to all Doses

Analysis of the genes whose expression was increased or decreased across time points for the same or different dose level showed only a small number of genes in common. A large number of the genes expressed were unique for given exposure time interval and dose level.

For instance, for genes whose expression increased (Figure 12a) at 0.01mg/g-diet (2-and 12-weeks), 3 genes were in common, and in case of 0.01mg/g-diet (2-weeks) and 0.1mg/g-diet (12-weeks), 8 genes were found to be common. Similarly, 0.01mg/g-diet at 2-weeks compared to 1mg/g-diet after 12-weeks (Figure 12b) showed 7 genes in common.

Furthermore, comparison of gene expression changes for genes whose expression increased (Figure 12c) between 0.1mg/g-diet (2-and 12-weeks), 4

genes were found to be in common. No genes were found to be (Figure 12c) in common when gene expression changes for 0.1mg/g-diet (2-weeks) increasers were compared to 1mg/g-diet (12-weeks) increasers.

On the other hand, comparison of gene expression changes for genes whose expression decreased (Figure 13a) between 0.01mg/g-diet (2-and 12-weeks) showed 4 genes in common, and only 2 genes in common between 0.01mg/g-diet (2-weeks) and 0.1mg/g-diet (12-weeks).

Also, comparison of 0.01mg/g-diet (2-weeks) and 1mg/g-diet (12-weeks) decreased genes (Figure 13b) showed 3 genes were in common, while only 1 gene was in common to all the doses in the time points compared.

In addition, comparison of gene expression changes for genes whose expression decreased (Figure 13c) between 0.1mg/g-diet (2-and 12-weeks) showed 1 gene in common, and comparing 0.1mg/g-diet (2-weeks) and 1mg/g-diet (12-weeks) showed 3 genes in common.

Other studies have also found uniquely expressed genes among classes of compounds and dose (Bartosiewicz et al, 2001; Waring et al, 2001). Dose-response studies of three chemicals (BaP, Dimethylnitrosamine, and CCl4) showed variation in gene expression patterns with dose (Bartosiewicz et al, 2001).

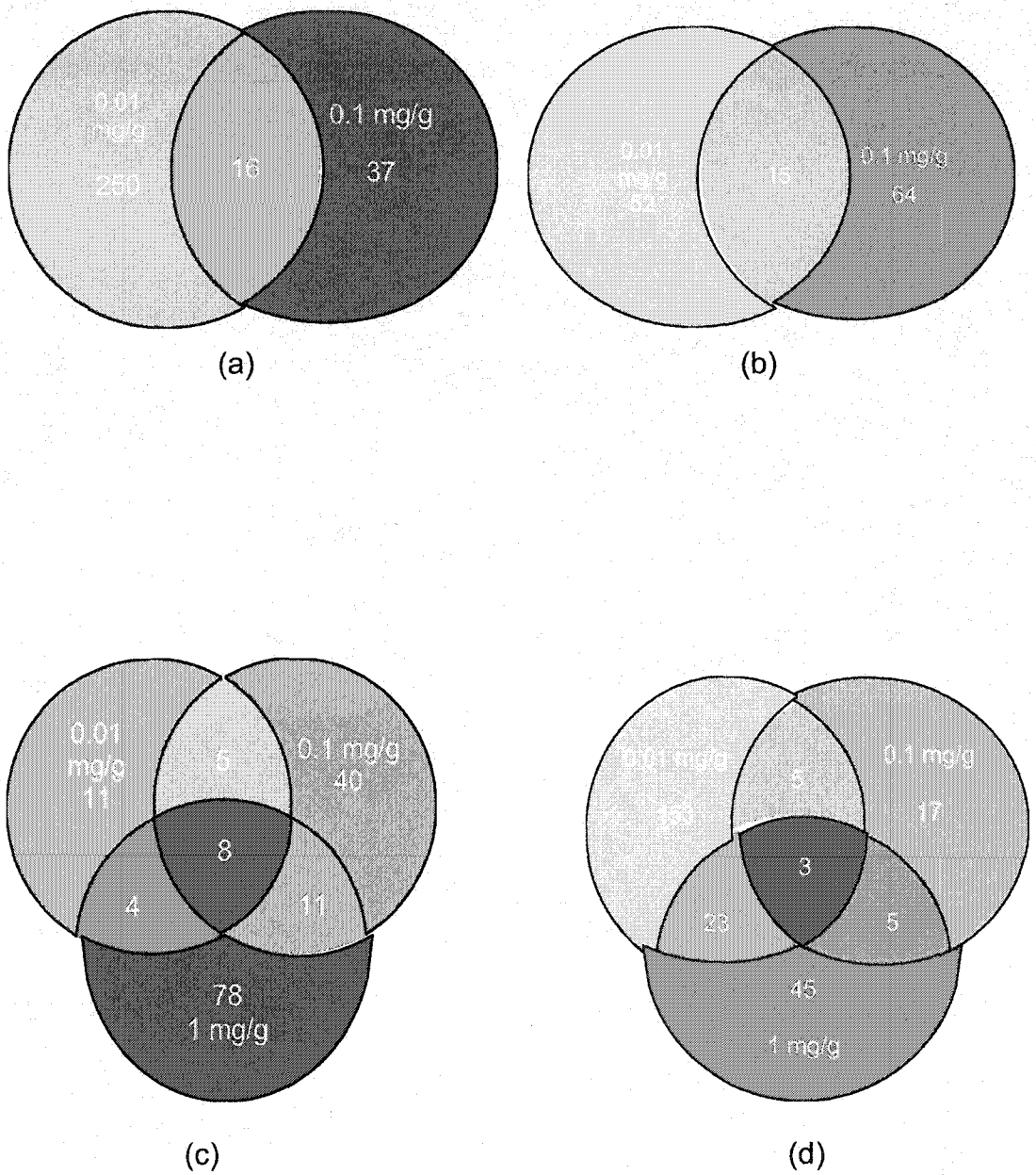
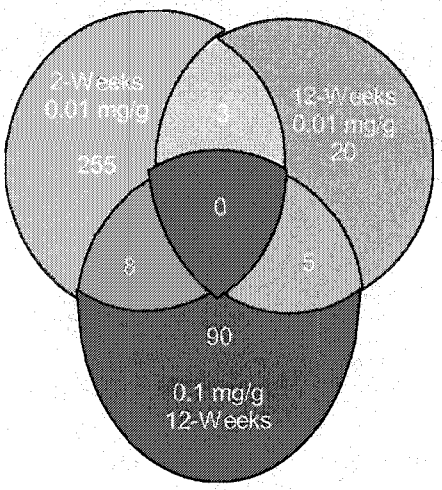
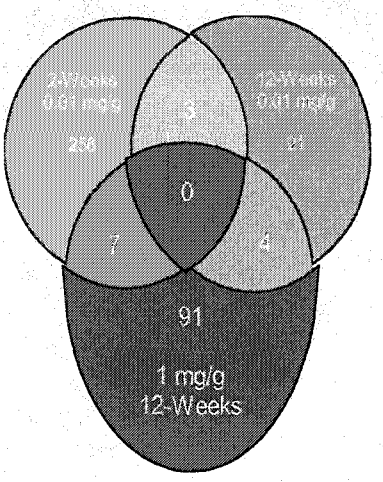


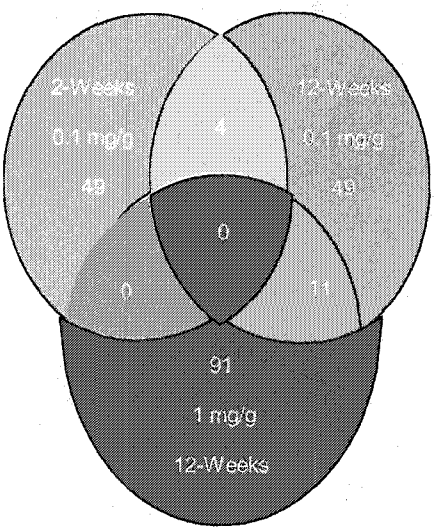
Figure 11. Venn Diagrams of Transcripts Unique or in Common to All Treatment Groups After 2 and 12-week Exposure to BaP. (a) Transcripts that Increased After 2-weeks, (b) Transcripts that Decreased After 2-weeks, (c) Transcripts that Increased After 12-weeks, and (d) Transcripts Decreased After 12-weeks of Exposure to BaP. Numbers in the intersection indicates genes in common between the two doses.



(a)



(b)



(c)

Figure 12. Venn Diagrams of Unique and Common Transcripts That Showed Increase Expression Across 2 and 12-week Dose Comparisons. (a) Dose Comparison for 0.01mg/g (2 and 12-weeks) and 0.1 mg/g (12-weeks). (b) Dose Comparisons for 0.01 mg/g (2 and 12-weeks) and 1.0 mg/g (12-weeks). (c) Dose Comparisons for 0.1 mg/g (2 and 12-weeks) and 1 mg/g (12-weeks). Numbers in the Intersection indicates genes in common for two or more dose and time points.

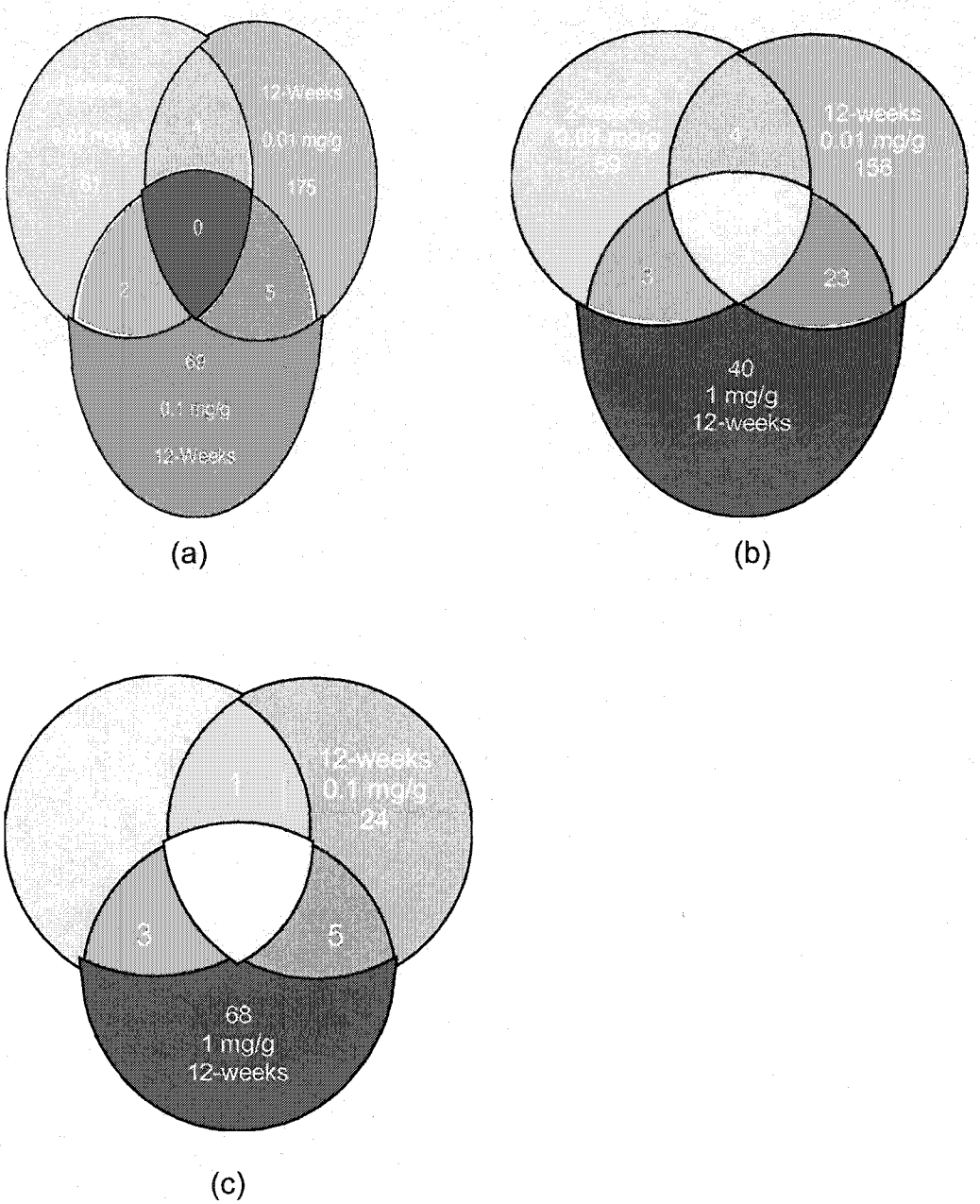


Figure 13. Venn Diagrams of Unique and Common Transcripts that Showed Decrease Expression Across 2 and 12-week Dose Comparisons. (a) Dose Comparison for 0.01mg/g (2 and 12-weeks) and 0.1 mg/g (12-weeks). (b) Dose Comparisons for 0.01 mg/g (2 and 12-weeks) and 1.0 mg/g (12-weeks). (c) Dose Comparisons for 0.1 mg/g (2 and 12-weeks) and 1 mg/g (12-weeks). Numbers in the Intersection indicates genes in common for two or more dose and time points.

Cluster and Principal Component Analysis of Genes

Affymetrix absolute files containing 8800 genes for the individual experiments were used to generate various clusters of genes. Because, it is possible to obtain different results depending on the clustering algorithm or method used, three separate clustering operations were performed; 1) Hierarchical Cluster, 2) Self Organizing Maps (SOM), and 3) Principal Component Analysis (PCA). Hierarchical cluster analysis as previously described in the methods section and by Eisen et al (1998) and Waring et al (2001) was applied across the various treatment groups using MAS 5.0 absolute files of individual experiments. In this analysis individual genes or transcripts are represented in rows, while time points are in columns. The average signal intensity of genes across similar treatments was first determined. This was normalized to the control signal of each treatment time point. The control normalized signal intensities ratios of each gene for all treatment were transformed to a $\log_2$ signal log ratio. These weighted signal log ratio were used to cluster each gene based on threefold change in signal log ratio (increase or decrease) from control gene expression. The hierarchical cluster showing similarities in the expression profile of genes is shown in Figure 14. The weighted average method (Eisen et al, 1998) was used to generate the hierarchical cluster using Euclidean distance as the similarity metric or measure. Euclidean distance measures how different two different profiles are taking into account the magnitude of the signals. Genes showing similarities in expression are joined together by the dendrogram on the left (Figure 14). The dendrogram

arrange the clustered genes in terms of their similarity to one another. The genes that are most similar are joined at low heights and those that are dissimilar are joined at larger heights. Individual cells are colored based on the log of fluorescent ratios; increasing ratios (>=3.0) indicative of increased expression over control are represented by decreasing gray intensity, while decreasing ratios (<=-3.0) are represented by increasing intensities of gray and reveal decreased gene expression. Gene annotations on the right (Figure 14) allow easy visualizations of groups of genes joined by a node on the dendrogram for biological relevant processes. Further resolution of the dendrograms is shown in Figure 15. Initial examination (Figure 14) of the data revealed apparent patterns of gene expression across experimental treatments; for example, a good number of genes are upregulated at the low dose as against the high dose after two-weeks of exposure to BaP. However, after 12-weeks of exposure, a significant number of genes were induced by the medium and high dose, with only few genes showing increase in the low dose treatment. These results (Figure 14) demonstrate the temporal separation of the gene expression response and most importantly, it confirms the idea of an adaptive versus the toxic response to toxicity over a given range of time. For instance, after two-weeks of exposure to the low dose, the expression of a large number of genes increased over the controls (Figure 14), however, after 12-weeks of exposure to the same dose, only few genes showed increased expression over the control level. This may be explained by the adaptive or tolerance response to sub-lethal concentrations of a toxic

compound (Neumann and Galvez, 2002; McDonald and Wood, 1993; Schmidt and Bradfield, 1996). McDonald and Wood (1993) demonstrated these adaptive responses in fish as occurring in three distinct physiological responses upon exposure to sub-lethal concentrations of waterborne metals. Initially, a shock phase that last for a few days is elicited, during which time the metal causes much of its morphological damage to fish gills. This shock phase is followed by a more gradual recovery phase when physiological conditions return back to normalcy or achieve a new steady state. These cellular turnover and repair process may cause increase tolerance to subsequent lethal exposures of the toxicant. On the other hand, as the dose (toxicity) increases as shown in Figure 14, the toxic response sets in and therefore an increase expression of genes was observed in both the medium and high dose groups. At the same time also, a good number of genes in medium and high dose groups after 12-weeks showed decreased expression, which may have been cause by the increase toxicity at the high dose or temporal regulation of these genes. The dendrogram on the left of the hierarchical cluster can be used to generate clusters of genes; 10 such clusters were generated as shown in Figure 16a. Principal component analysis of the 10 clusters (Figure 16b) was done to visualize the distribution of the genes in space and also determine to which clusters carry genes with the most variability. A 3-dimensional principal component analysis plot is shown in Figure 16b. Genes in cluster 8 and 10 showed strong downward shift and in addition some genes in cluster 4 and 9 and also show a negative

shift in their variability. On the other hand, genes in cluster 3, 6 and 7 show a positive shift in their variability. The use of this analysis after hierarchical clustering allows transformation of the data to a smaller dimension that shows genes with the most variability in the data.

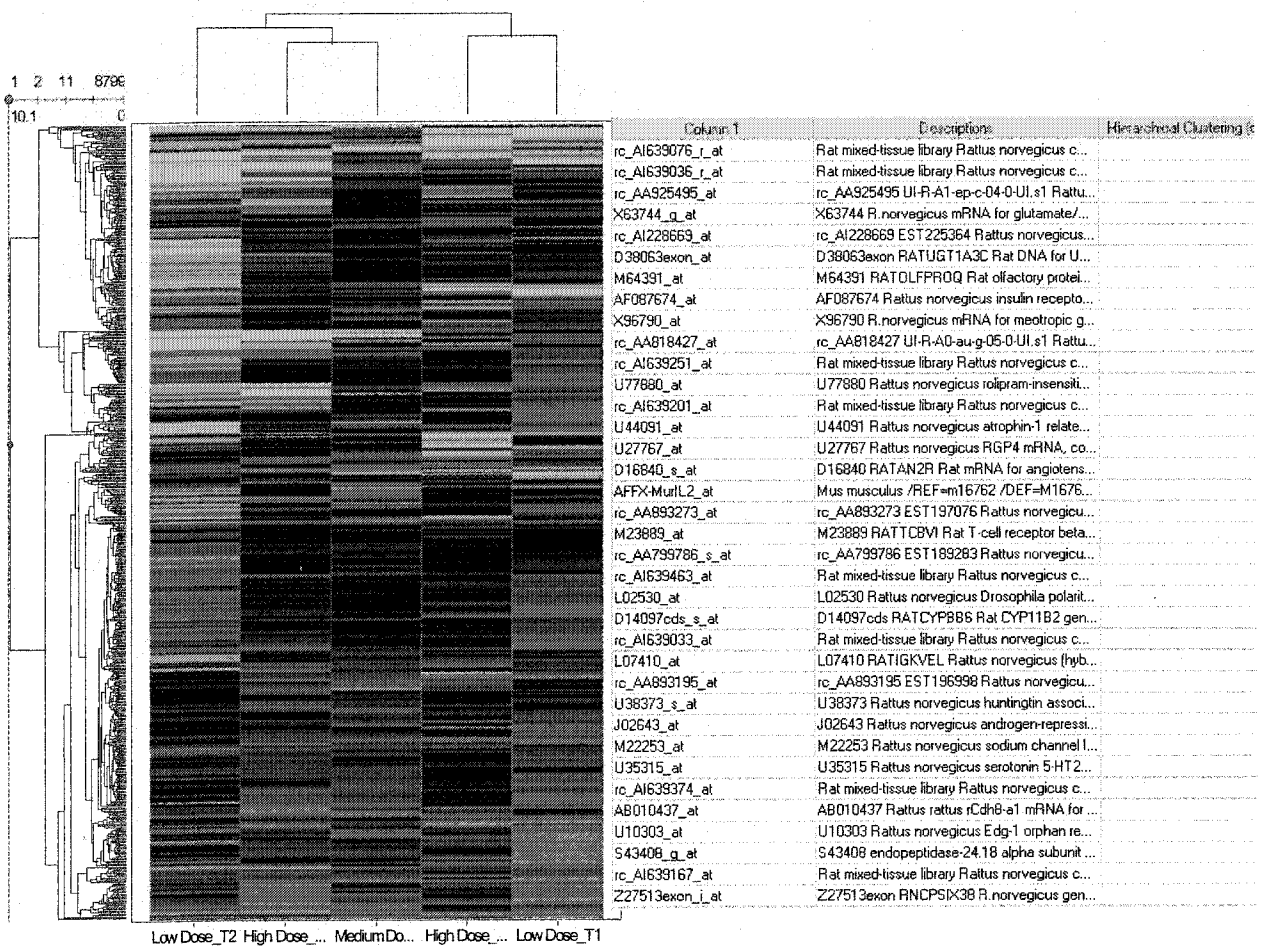


Figure 14. Hierarchical Cluster of Genes Across Treatment Groups. The dendrogram on left connects genes with similar expression pattern. Partial annotation of the genes is shown on the right. A three-fold increase or decrease in expression indicated by decreasing or increasing intensities of grayness, respectively.

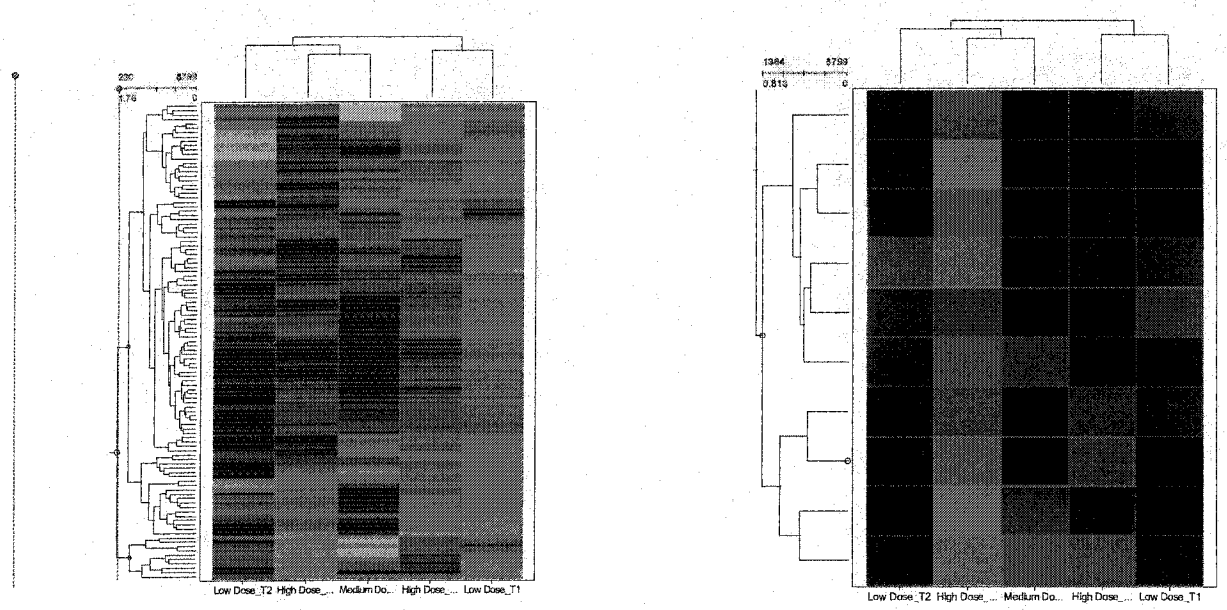
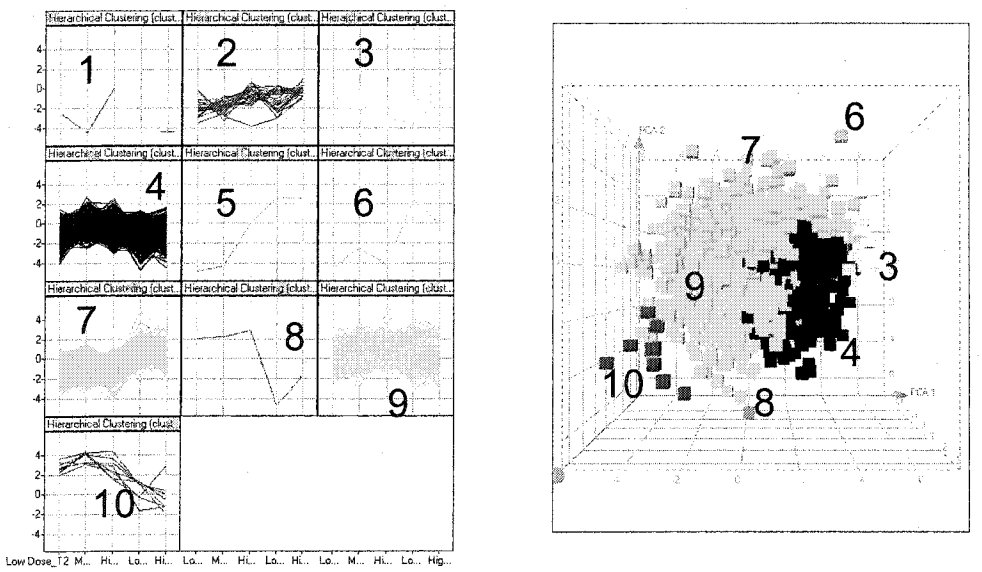


Figure 15. Zoom Version of a Hierarchical Cluster Node.

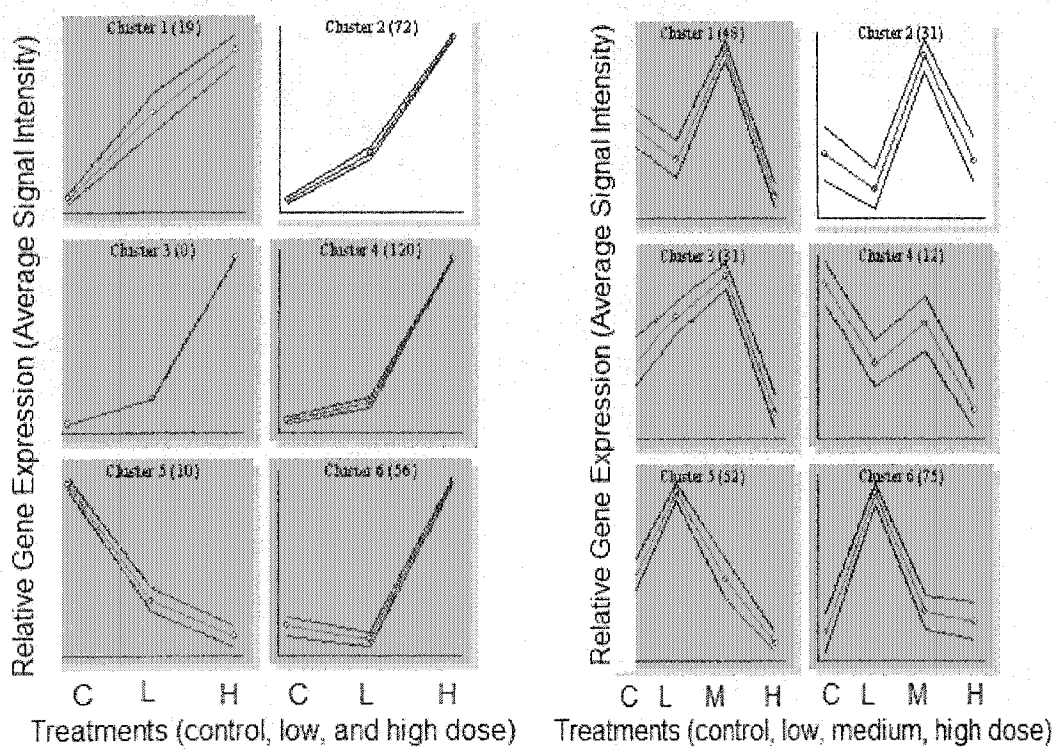


(a) (b)

Figure 16. Principal Component Analysis of Hierarchical Clusters. (a) Ten Clusters extracted from the Hierarchical Cluster. (b) Principal Component Analysis of the ten clusters.

The results from the SOM analysis are presented in Figure 17. The SOM program in the Affymetrix data mining tool analysis function was used for clustering the dataset. SOM is well suited for exploratory data analysis, to expose the fundamental patterns in the data. The underlying structure can be readily explored by varying the geometry of the node. The nodal geometry was set to identify six clusters as described in materials and methods section. Each of the clusters represent a group of genes that peak at a particular dose and time and have a similar pattern, therefore each SOM cluster represents the average pattern of genes within the cluster. The y-axis represents relative expression levels (average signal intensity) and the x-axis represents the treatments groups and control.

For the 2-weeks SOM cluster of genes (Figure 17a) showed 19 genes clustered in cluster 1, 72 genes in cluster 2, 0 genes in cluster 3, 120 genes in cluster 4, 10 genes in cluster 5, and 56 genes in cluster 6. After 12-weeks of exposure, the SOM cluster of genes (Figure 17b) showed 48 genes in cluster 1, 31 genes in cluster 2, 31 genes in cluster 3, 12 genes in cluster 4, 52 genes in cluster 5, and 75 genes in cluster 6. Table 10 and 11 shows summaries of the genes present in SOM clusters of the 2-and 12-weeks exposure period. As with hierarchical clustering, the clusters represent distinct patterns of detectable gene expression. Affymetrix Accession numbers and associated GenBank ID's are given.



(a) (b)

Figure 17. Self Organizing Maps (SOM) Clusters of Genes Across Treatment Groups. (a) Six SOM clusters of 2-week BaP differentially expressed genes. (b) Six SOM clusters of 12-week BaP differentially expressed genes.

After 2-weeks of exposure, the results of the clusters (Figure 17a) show some distinct patterns of gene expression. For example, cluster 1, consisted of genes in their exponential phase of induction, where as cluster 5 showed genes that were initially induced and showing decrease in expression over time. In addition, clusters 2, 3, 4, 6 grouped together genes that showed late induction from an initial lag. The gene names and Gene Ontology molecular function description of the annotated genes from each of the SOM clusters after 2-weeks exposure to BaP are shown in Appendix A1 to A5. The list of

annotated genes from the SOM cluster 1 (Appendix A1) are involved in molecular functions such as metabolism (aldehyde dehydogenase, alcohol dehydrogenase, steroid 5 alpha-reductase 1, cytochrome p450 II C6), peroxisome organization (peroxisomal membrane protein), and other functions. In cluster 2 (Appendix A2), the annotated genes show a more diverse molecular function; enzymatic catalysis (acyl-coA synthetase, argininosuccinate lyase, ATP citrate lyase, glycine-N-methyltransferase), immune and inflammatory response (interleukin-1 beta, intercellular adhesion molecule 1, fibrinogen), chaperone activity (DnaJ homolog), RAS oncogenes (RAB11a), and so on. The genes with diverse molecular function are also clustered together in cluster 4 (Appendix A3). These include genes involved in enzymatic catalysis (palmitoyl-protein thioesterase 2, hepatic lipase, 3-hydroxy-3-methylglutaryl coA lyase), transcription regulators and kinases (pyruvate kinase, *NF-E2* related factor 2, adenylate kinase), stress and defense response (clusterin, *TGFB* inducible early growth response, *CD81* antigen, leukotriene B 4, 12-hydroxydehydrogenase), hormone activity, signal transduction, transporters. A much smaller subset of genes consisting mostly of protein kinases (platelet derived growth factor receptor, *SH3*-domain kinase), phosphatases, and calcium ion binding proteins (S100 calcium-binding protein, matrix *Gla* protein), *GTP* binding protein such as tubulin are represented in the cluster 5 (Appendix A4). Also, cluster 6 (Appendix 5) consist of various transcription factors (forkhead box A2, one cut domain family member 1, dual specificity phosphatases), apoptosis (programmed cell

death 8), ATpases, cell adhesion, enzymatic metabolism, and other molecular functions. On the other hand, after 12-weeks of exposure (Figure 17b), the expression patterns of these genes differ from the 2-weeks genes. For instance, cluster 6 showed genes that were initially induced, followed by a peak and then gradual decrease in expression to a more stable level. Similar patterns are shown by genes in cluster 3 and 5. In the case of cluster 1, 2, and 4, the expression of genes initially decreased, followed by a gradual increase to peak levels and subsequently decreased in their expression. The names and Gene Ontology molecular function description of the annotated genes from each of the SOM clusters after 12-weeks of exposure to BaP are shown in Appendix B1 through B6.

As shown in Appendix B1, cluster 1 consists of various growth factors (insulin-like growth factor, activin beta E), apoptosis and cell cycle regulators (similar to cytochrome c, B-cell lymphoma 2), DNA binding, metal ion binding and transcription factor such as ATpase Calcium transporter, carbonic anhydrase 3, nuclear factor I/A, one cut domain, and cold shock domain protein A. Genes in cluster 2 (Appendix B2) consist mostly of electron and cation transporters (solute carrier family 26, ATpase Hydrogen transporter, ceruloplasmin, cytochrome b, solute carrier family 28), transcription regulators (mitogen activated protein kinase, integrin alpha 1, cofilin 1), metabolism (apolipoprotein B, steroid 5 alpha reductase 1), and others.

In cluster 3 (Appendix B3) of the 12-weeks SOM, the genes represented include the DNA repair and DNA binding (*DnaJ Hsp40* homolog, zinc finger

protein 354A, hairy and enhancer of split 1, nuclear receptor subfamily, *CD151* antigen, and CCAAT/enhancer binding protein), and transferases such as glutathione-s-transferase a2, catechol-o-methyltransferase, and dopa/tyrosine sulfotransferases. Also, cluster 4 (Appendix B4) consists of nuclear receptor subfamily, cytochrome oxidase subunit VIc, pyruvate kinase, and UDP glycosyltransferase 1 polypeptides family of genes. In addition, cluster 5 (Appendix B5) of various detoxification and metabolism genes (NADPH cytochrome p450 reductase, Cytochrome p450 1a2, ATP citrate lyase, UDP glucuronosyltransferase 1 polypeptide family, aldo-keto reductase family, sterol-C4-methyl oxidase, pyruvate dehydrogenase kinase), *CD74* antigen, and rat senescence marker. Similarly, cluster 6 consists of detoxification and xenobiotic metabolism genes (aldehyde dehydrogenase, squalene expoxidase, cytochrome p450 1a1, NADPH dehydrogenase quinone 1, cytochrome p450 1a2, glutathione-s-transferases), oncogenes and GTpases (*RAP1B* and *RAB7 ras* ooncogene families), and other genes.

The data from the SOM clusters is comparable to some of the panels from hierarchical clusters. Recent studies (Tamayo et al, 1999; Burton et al, 2002; Toronen et al, 1999; Waring et al, 2001; Amin et al, 2004; Eisen et al, 1998) have also employed hierarchical and SOM clustering to identify gene expression patterns and determine mechanisms of toxicity. For example, Waring et al. (2001) used hierarchical clustering to classify 15 hepatotoxins based on mechanisms of toxicity. Recently, Amin et al (2004) has used hierarchical clustering and PCA to identify putative gene based markers of

renal toxicity in male Sprague-Dawley rats. Similarly, Tamayo et al. (1999) and Burton et al. (2002) has also applied SOM clustering to interpret patterns of gene expression. Principal Component Analysis (PCA) was also applied to data as previously discussed in the methods section. The results (Figure 18a) of the PCA is shown as a 3D scatter plot with three PC; PCA1, PCA 2, and PCA3. A large percentage (>50 %) of the variability in the expression of the genes was associated with PCA 1, while PCA2 and PCA3 accounts for the remaining variability. Time 1 (2-weeks) and Time 2 (12-weeks) are indicated by blue and red dots (Figure 18a), respectively.

As in hierarchical clustering described above, the control normalized signal log ratios for treatment group at each time point from the Affymetrix MAS 5.0 software were loaded into Spotfire Decision site for functional genomics, and PCA was conducted. The PCA scatter plot of the genes (Figure 18a) had a similar pattern when compared to the PCA plot generated from the hierarchical cluster (Figure 16b). The PCA1 and PCA2 cluster of genes were considered the most relevant in terms of providing useful gene expression information and similarities in terms of their variation with dose. The PCA plot of the 2-weeks and 12-weeks showed similarities in the distribution of the genes in 3-D space. The distribution of the genes in the two time points is comparable to two parallel planes with only difference being a vector; with a negative direction. If one were looking at the data from a 2D point of view, the only difference rests on PCA1, which carries most of the variability. The use of PCA for multidimensional visualization of gene expression in order to

observe the spatial distribution of the treatment groups is extensively discussed in the literature (Yeung and Russo, 2001; Parmigiani et al, 2003). In a related study, Amin et al (2004) used PCA to profile the dose-response and temporal regulation of genes by two classes of toxicants. This approach offers opportunity to visualize expression patterns that can reflect similarities in biological responses. The visualization of high dimensional data in two-or-three dimensional components reveals unsupervised clusters within the data. The result of the PCA as shown in Figure 19 and Table 9 can be scored, and based on those PCA scores, unknown samples can be determine and placed into one of the previously identified clusters where gene expression is clearly linked to biological processes.

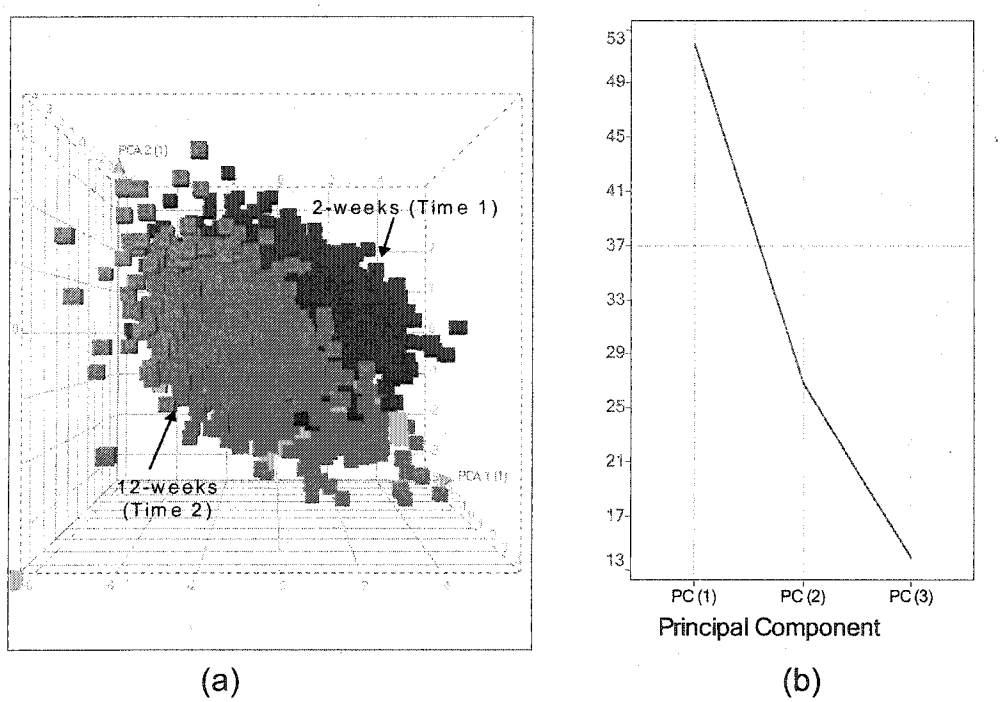


Figure 18. Principal Component Analysis of Gene Expression Across Time. (a) 3-Dimensional PCA Scatter Plot showing variation in Gene Expression for the 2 and 12-week time point. (b) Percent variability plot of the PC1, PC2, and PC3 variables

Table 9

Eigenvalue Loadings for PCA1, PCA2, PCA3 Variables

Principal Component	Time	Low Dose	Medium Dose	High Dose
PC (1)	0.153	0.873	0.403	0.226
PC (2)	0.539	0.318	-0.691	-0.362
PC (3)	-3.5e-2	2.8e-2	0.452	-0.891

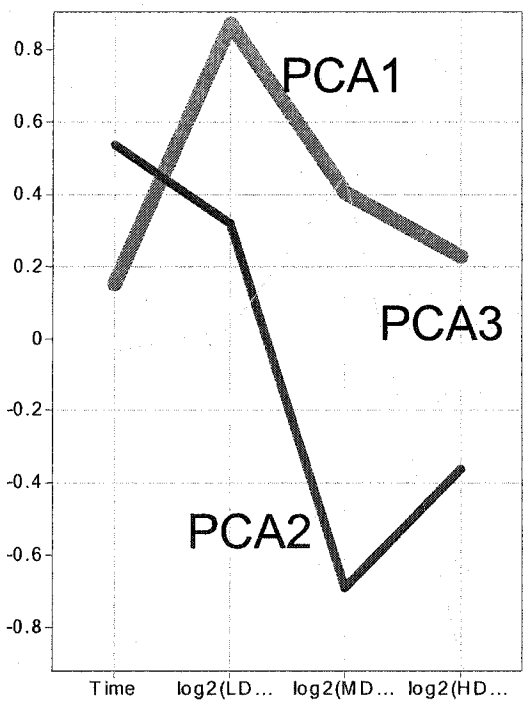


Figure 19. Scree Plot of PCA1, PCA2, and PCA3 Variables. Scree plots show variability in the gene expression profiles with time and dose.

Table 10

Affymetrix Accession Numbers of Representative Genes or Transcripts Present in the Six SOM Clusters of the 12-weeks Exposure Period

Cluster 1	Cluster 2	Cluster 3	Cluster 5	Cluster 6
AA799276_at	AA801441_at	AB017544_g_at	AF014503_at	AB002558_at
AF030091UTR#1_g_at	AB009890_at	AFFX-BioB-M_at	AF045464_s_at	AF001898_at
AF037072_at	AF034582_g_at	AFFX-BioC-3_at	AJ011656cds_s_at	AF003835_at
AF063102_at	AFFX-BioB-3_at	AFFX-BioC-5_at	D28339_s_at	AF036761_g_at
AF077000_at	D12769_at	AFFX-BioDn-3_at	D83796_s_at	AF061971_at
AF079864_at	J05035_g_at	AFFX-BioDn-5_at	E01184cds_s_at	AF062389_at
AF089825_at	L23413_at	AFFX-CreX-3_at	E01524cds_s_at	AF072411_at
AFFX-BioB-5_at	M13100cds#2_s_at	AFFX-CreX-5_at	E12625cds_at	AF072411_g_at
D00698_s_at	M13100cds#3_f_at	D13417_g_at	J03179_at	D10655_at
D28557_s_at	M13100cds#6_f_at	D86580_at	J03179_g_at	D10655_g_at
J00738_s_at	M27440_at	D86745cds_s_at	J05210_at	D30040_at
J03865mRNA_f_at	M57718mRNA_s_at	D88890_at	J05210_g_at	D37920_at
L38482_at	M58758_g_at	D89375_s_at	K01934mRNA#2_at	E00717UTR#1_s_at
M91235_f_at	M91599mRNA_g_at	J03190_at	K02815_s_at	E00778cds_s_at
rc_AA799538_g_at	rc_AA817854_s_at	J03190_g_at	L00320cds_f_at	E03358cds_at
rc_AA800029_at	rc_AA852046_s_at	M58587_at	L38615_g_at	J02679_s_at
rc_AA800787_at	rc_AA859612_f_at	M93257_s_at	M10068mRNA_s_at	J05460_s_at
rc_AA859848_at	rc_AA859966_i_at	M96548_at	M21476_s_at	K03241cds_s_at
rc_AA859897_at	rc_AA894207_g_at	rc_AA800005_at	M25157mRNA_i_at	L12380_at
rc_AA859994_at	rc_AI105348_i_at	rc_AA892775_at	M26127_s_at	L13619_at
rc_AA860014_i_at	rc_AI171355_s_at	rc_AA893239_at	M74067_at	L22294_at
rc_AA874926_at	rc_AI171630_s_at	rc_AA893384_g_at	M76767_s_at	M26594_at
rc_AA875097_at	rc_AI180108_at	rc_AA943892_at	M89945mRNA_at	M33329_f_at
rc_AA892522_at	rc_AI639043_at	rc_AA946040_at	M89945mRNA_g_at	M36151cds_s_at
rc_AA893664_at	rc_AI639108_at	rc_AI010632_s_at	rc_AA799778_at	M37584_at
rc_AI008815_s_at	rc_AI639120_at	rc_AI011998_at	rc_AA852004_s_at	M63983_s_at
rc_AI071866_s_at	U58858_at	rc_AI171085_at	rc_AA892828_at	rc_AA799279_g_at
rc_AI176460_s_at	U66723_s_at	rc_AI175935_at	rc_AA945050_f_at	rc_AA799326_s_at
		Cluster 4		
		J02612mRNA_s_at		
		L21711_s_at		
		M25804_at		
		M25804_g_at		
		M27467_at		
		M33025_s_at		
		rc_AA859372_s_at		
		rc_AA945573_f_at		
		X05684_at		

Table 11

Affymetrix Accession Numbers of Representative Genes or Transcripts Present in the Six SOM Clusters of the 2-weeks Exposure Period

Cluster 1	Cluster2	Cluster 4	Cluster 5	Cluster 6
AB002558_at	AA685903_at	AB000491_at	AF023087_s_at	D12769_at
AF025308_f_at	AB012933_at	AB005547_at	rc_AA799448_g_at	D12769_g_at
D13667cds_s_at	AB012933_g_at	AF012714_at	rc_AA818593_g_at	D32249_s_at
D14989_f_at	AF023302_s_at	AF029240_at	rc_AA892498_at	D89655_at
E03358cds_at	AFFX_rat_GAPDH	AF061971_at	rc_AI012030_at	J03190_g_at
E12286cds_at	AJ223355_at	AF089825_at	rc_AI232379_at	J03863_at
M14776_f_at	D00913_g_at	AJ000347_g_at	rc_AI639058_s_at	L00191cds#1_s_at
M21476_s_at	D13921_s_at	D00569_at	U90261UTR#1_g_at	L09647_at
rc_AA893242_g_at	D13978_s_at	D00698_s_at	V01227_s_at	L28135_at
rc_AA925473_g_at	D37934_g_at	D00729_g_at	X06916_at	L36532_s_at
rc_AA996484_g_at	D85435_at	D00753_at		M77479_at
rc_AI102838_s_at	D85435_g_at	D10874_at		M91235_f_at
rc_AI171734_s_at	D87839_g_at	D10874_g_at		S81478_s_at
S35751_f_at	D87991_at	D13623_g_at		U02553cds_s_at
S81448_s_at	D90401_at	D13907_g_at		U06230_s_at
X13527cds_s_at	D90401_g_at	D21215cds_s_at		U18762_at
X53054_g_at	E01050cds_g_at	D30647_at		U19614_g_at
X70223_at	E03859cds_s_at	D30649mRNA_s_at		U32314_g_at
X72792cds_s_at	E13557cds_s_at	D45254_g_at		X13119cds_s_at
	J00735_at	D88250_at		X55298_at
	J05210_at	D90211_s_at		X78997_at
	K00750exon#2-3_at	D90265_s_at		X95096_at
	K02817cds_s_at	D90404_at		Y14933mRNA_s_at
	M15481_g_at	D90404_g_at		
	M17412_at	E01184cds_s_at		
	M26247_at	J00728cds_f_at		
	M57718mRNA_s_at	J04171_at		
	M62832_at	J05035_at		
	M64755_at	J05035_g_at		

Influence of BaP on Gene Expression Profiles and Biological Processes

Table in Appendix C shows a summary of the annotated genes that were regulated and found to be relevant in the BaP-Ah receptor mediated as well as other signaling pathways leading to cancer and other health effects. The effects of BaP on the Ah receptor mediated pathway as deduced from the microarray data is shown in Figure 22. A large number of genes were

differentially expressed in response to varying doses of BaP. In addition, the results show temporal regulation of gene expression of a good number of genes. The gene expression profile of our microarray data (Figure 20) show variations in gene transcription patterns, under the influence of temporal and dose-response relationship. A good number of the genes show greater than three-fold changes in their expression from the control and vary between treatments groups.

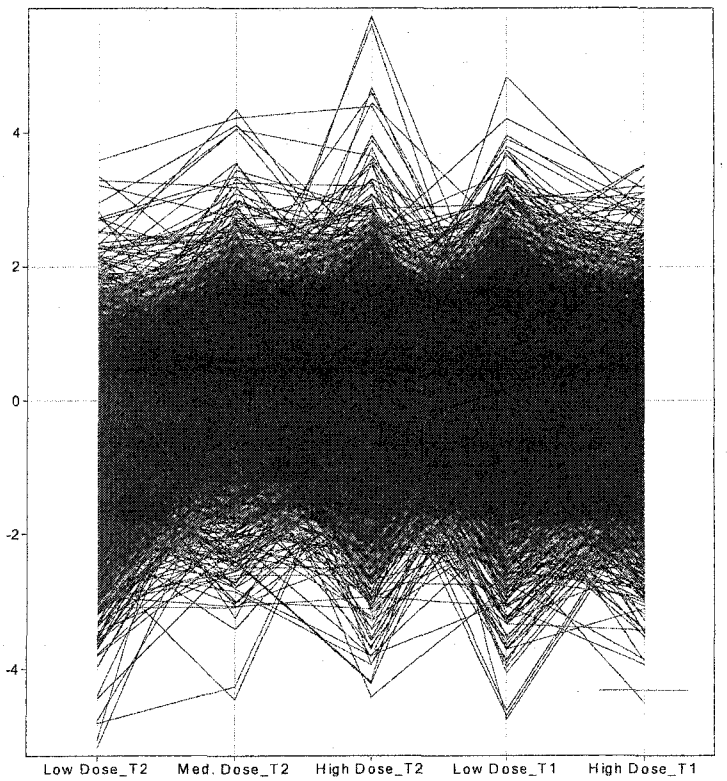


Figure 20. Gene Expression Profiles of Genes Across Treatment Groups and Time.

Both dose and exposure time interval was found to be important factors in determining not only the number of genes that are regulated, but also the time period for which these genes are increased or decreased (Appendix C). For instance, several genes in our analysis were up-regulated after 2-weeks but their expression decreased after 12-weeks of exposure (Appendix C) to control levels or below. Similarly, within exposure time intervals, variation in expression of some genes with dose (Appendix C) was observed. Some genes tend to be induced or repressed at high, medium or low doses in a dose-dependent pattern, while others tend to be induced or repressed only at specific dose level at each exposure time interval. This apparent non-linear dose-response relationship has been noted for BaP and other compounds in previous studies (Bartosiewicz et al, 2001).

Some of the genes identified in the microarray data and differentially expressed (threefold or greater) in at least one of our arrays (sample) were classified into biologically relevant pathways, based on information from Affymetrix NetAffix, NCBI databases, GenMAPP, Gene Ontology (GO) or Ingenuity Pathway knowledgebase. Several recent validation studies (Amin et al, 2004; Ezendam et al, 2004) consider one present call in at least one of their samples to determine which genes were differentially expressed. Also, Affymetrix absolute calls (Absent to absent, Absent to Marginal, Absent to present, present to absent, present to present, present to marginal) were utilized to select other genes as differentially expressed. After two weeks of exposure to 0.01 and 0.1 mg/g of BaP in the diet, a good number of the

genes induced were transcription regulators, anti-apoptotic genes, apoptotic genes, and genes involved in metabolic pathways, oxidative stress, cell cycle regulators and calcium metabolism (Appendix C). Subsequently, the induction of a large number of these genes was repressed or decreased over time. For instance, some of the genes upregulated at this time point are *p53, p21, Bax* (apoptotic genes), *Bcl-2, Bcl-2 like*, (anti-apoptotic), *c-Ha-ras*, *ras-related p23* (Oncogenes), antioxidants *GST-yc2, GST-ya*, peroxidases, NADPH reductases, aldehyde dehydrogenase-3A1 (Acute phase proteins), and others. A representing pathway map showing the induction and repression of apoptotic and anti-apoptotic genes is shown in Figure 21.

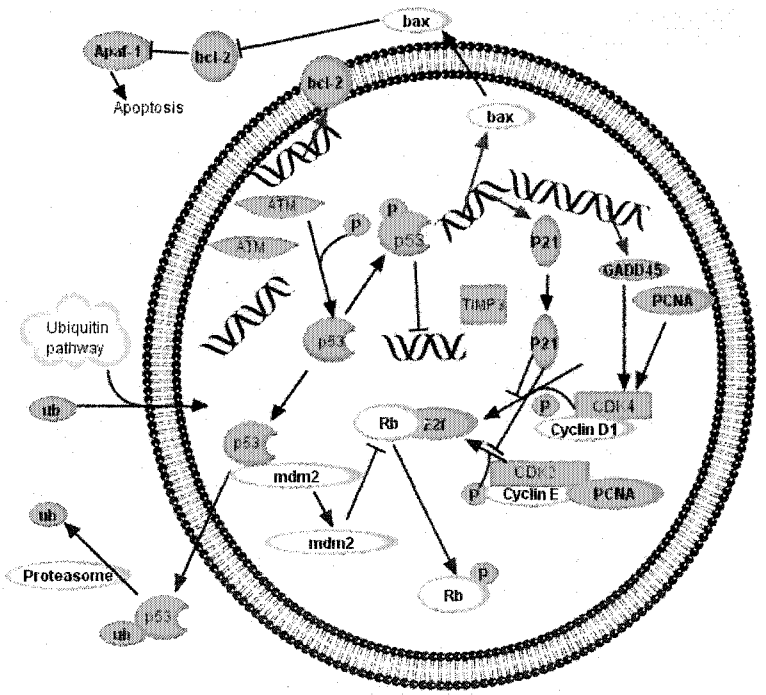


Figure 21. Apoptosis Pathway Genes Regulated by BaP. Transcripts induced by BaP are shaded gray.

After 12-weeks of exposure, the gene expression profile consisted of detoxification enzymes (Cytochrome p450), some heat shock proteins and DNA repair enzymes (Appendix C). This apparent temporal phase distinction in gene expression has been observed in other studies of other compounds (Yoneda et al, 2001; Huang et al, 2001). Using DNA microarrays, Yoneda et al. (2001) observed that gene expression in human epithelial cells undergo distinct temporal phases when exposed to oxidants such as hydrogen peroxide and smoke.

Previous studies (Parish et al, 1998; Binkova et al, 2000; Ramet et al, 1995; Mounho et al, 1997) on the effects of BaP on expression of some of these genes confirms this finding. For example, Ramet et al. (1995) show that *p53* protein expression is correlated with BaP-DNA adducts in carcinoma cell lines. Similarly, Salas and Burchiel (1998) have demonstrated an association between BaP, increase in intracellular calcium and induction of Apoptosis in Daudi cells.

In addition, the microarray data also reveals that BaP may be capable of disrupting key rate limiting transcripts or enzymes in various metabolic pathways as shown in Figure 23 and cause health effects. For instance, the results show disruption of the steroid hormone signaling via up regulation of the rate limiting enzyme of cholesterol (HMG-CoA reductase), suppression of *Cyp450C17* hydroxylase, and by suppression of Androgen receptor (AR) may affect estrogen/androgen metabolism, which leads to endocrine disruption. Studies by Vinggaard et al. (2000) support the microarray finding that BaP

affect androgen receptor activation. Using Chinese Hamster Ovary cells (CHO), Vinggaard et al (2000) show that BaP act as antagonists of the AR. Additionally, differential expression of cytochrome p4501 gene family (*Cyp1a1, Cyp1a2*), *Gst*, and other genes involved in detoxification was observed in at least one of the doses in this study. BaP is known to induce these genes through binding to and activation of the Ah receptor (Barouki and Morel, 2001; Tian et al, 1999; Denison and Heath-Pagliuso, 1998; Schmidt and Bradfield, 1996). This is considered to be one of the major biochemical pathways through which BaP is metabolically activated to cause cancer and other health effects. The exact mechanism by which BaP causes cancer and other toxic health effects is not fully understood. As shown in Figure 22, binding to the aryl hydrocarbon Ah receptor and the subsequent metabolic activation and transport of the receptor from the cytosol to the nucleus is linked to the increase transcription of *Cyp1a1, Gst*, and several other genes containing the dioxin response elements (DRE) promoter. Increase transcription of *Cyp1a1* genes leads to increase in *Cyp1a1* enzyme activity and oxidative conversion of BaP to reactive intermediates capable of binding to cellular macromolecules such as DNA; this induces mutations and/or oxidative DNA damage. The resulting DNA adducts (Canova et al, 1998; Chakravarti et al, 1995; Cheh et al, 1993; Miller and Miller, 1976) formed are considered to be the first step in the initiation of cancers and other health effects in mammals. Also, the induction of *Gst-mu* in mice administered BaP intraperitoneally has been observed by Bartosiewicz et al (2001) using a

limited set of cDNA arrays. Gst are involved in detoxification mechanisms and their up-regulation may due to oxidative stress caused by BaP metabolism in the cell. These results are also consistent with other studies on the oxidative alteration of DNA and protein, and antioxidant enzymes in several organs (liver, kidney, lungs) of rats treated with BaP (Bartosiewicz et al, 2001; Kim et al, 1997)

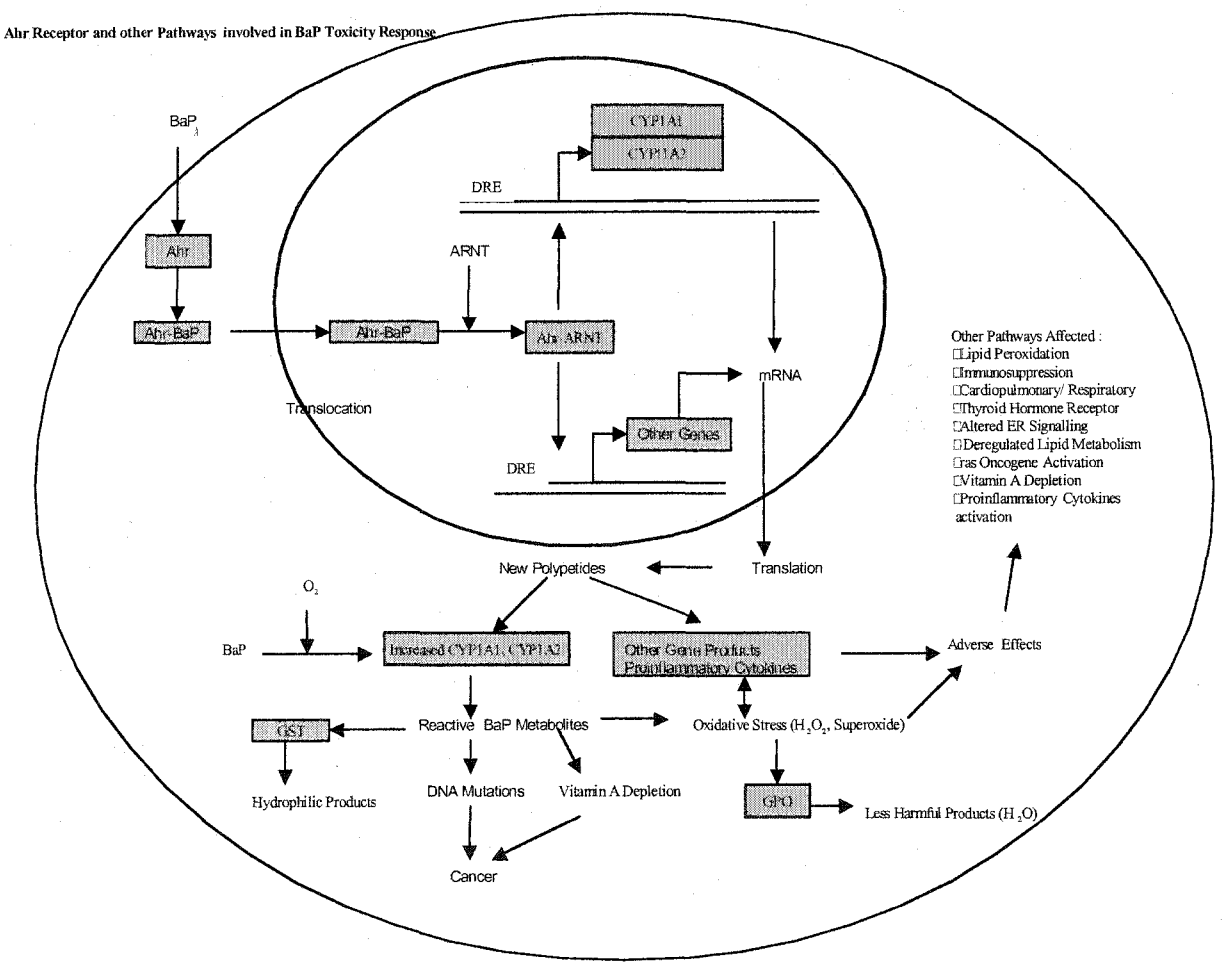


Figure 22. Ah Receptor Mediated Pathways of BaP Activation

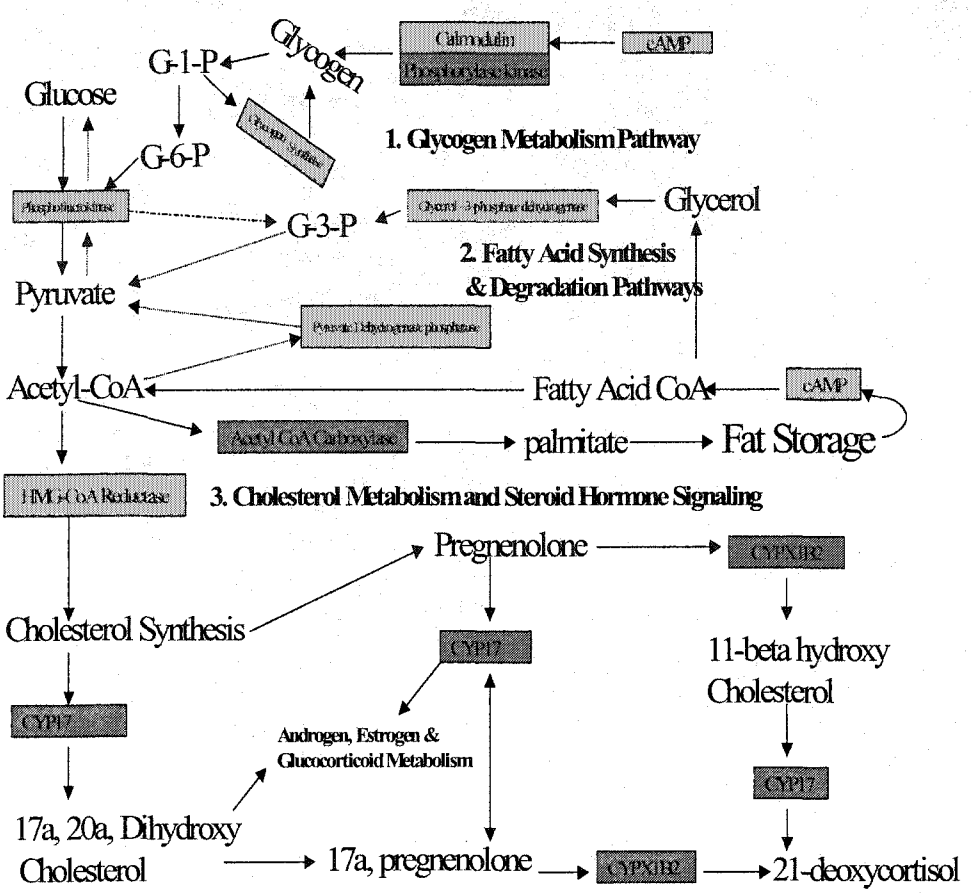


Figure 23. Metabolic and Steroid Hormone Pathways Regulated by BaP. Increase or Decrease Expression of Transcripts is indicated by light gray or dark gray, respectively.

Ingenuity Pathway Analysis of the Microarray Data

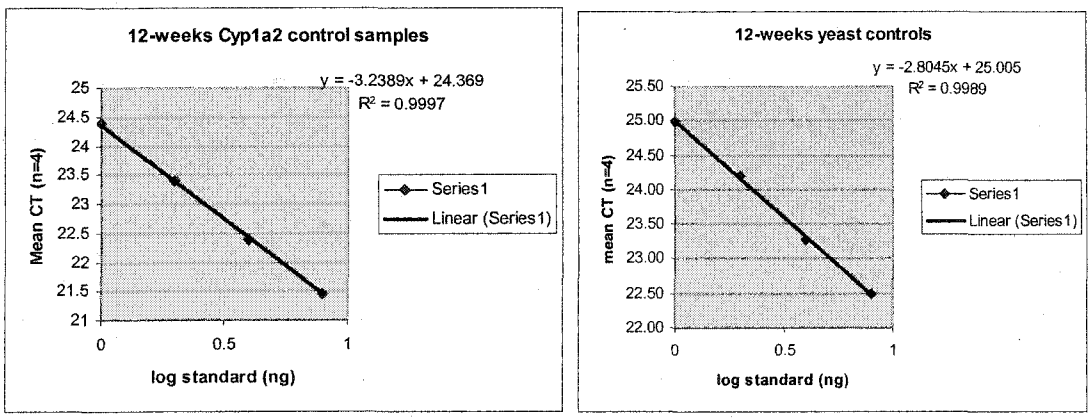
As previously described in the materials section, Ingenuity Pathway Analysis was used to construct biological pathways based on the genes that were differentially expressed threefold or greater by BaP. Affymetrix microarray data was uploaded into the Ingenuity knowledgebase system and program was run to include as many nodes. The results of the pathway analysis

suggest that BaP may have a wide ranging effect on a host of pathways (Appendix D1). For instance, BaP may regulate *IL-1 beta* activity via the *p38 MAPK* pathways and *NFkB* pathway shown in Appendix D2. Also, observed are the signal transduction involved in the regulation of *Bcl-2* (Appendix D3) and the tumor suppressor gene *p53* (Appendix D4), which plays a central role in cell cycle regulation, apoptosis, and cell death.

Real-Time Quantitative PCR (RT-QPCR) Analysis

Based on the results of the microarray data, we selected six genes for further analysis and for validation of our GeneChip data. The candidate genes selected were based on their association with health effects and involvement in pathways relevant to the understanding of the mechanisms BaP toxicity. As previously discussed, we used real-time quantitative one-step Taqman PCR to analyze our target genes of interest. The results with minimal variations confirm our findings from the microarray experiments. Some of the variations are explained by the differences in sensitivity and resolution between the GeneChip and RT-QPCR; higher fold changes are usually observed with RT-QPCR than by microarray. In order to control for variations due to mRNA isolation or RT-QPCR conditions, 50 ng of yeast mRNA was spiked into each sample during mRNA extraction. Also, RT-QPCR for each gene of interest was run in conjunction with yeast actin for each sample to normalize the expression levels.

The mRNA levels of each gene were quantified using Taqman ABI PRISM 7700 Sequence Detection System (Applied Biosystems). Standard curves for both the yeast internal control and the target genes was generated from serial dilutions of untreated control samples, and used to determine the amount of mRNA for each gene and yeast in all experimental samples. A representative standard curve measurement for *Cyp1a2* and yeast actin internal control gene is shown Figure 24 below. The relative expression level of each gene was computed with respect to yeast mRNA for each treatment group to account for variation in quality of mRNA or RT-PCR conditions. To control for genomic DNA amplification, primers were constructed to include one intron, and in addition, no amplification control curves were ran alongside each gene's standard curve.



(a) (b)

Figure 24. Standard Curves from Serial Dilutions of Untreated Control Samples. (a) serial dilutions (log standard) for Cyp1a2 gene standard curve. (b) serial dilutions (log standard) for yeast actin internal control standard.

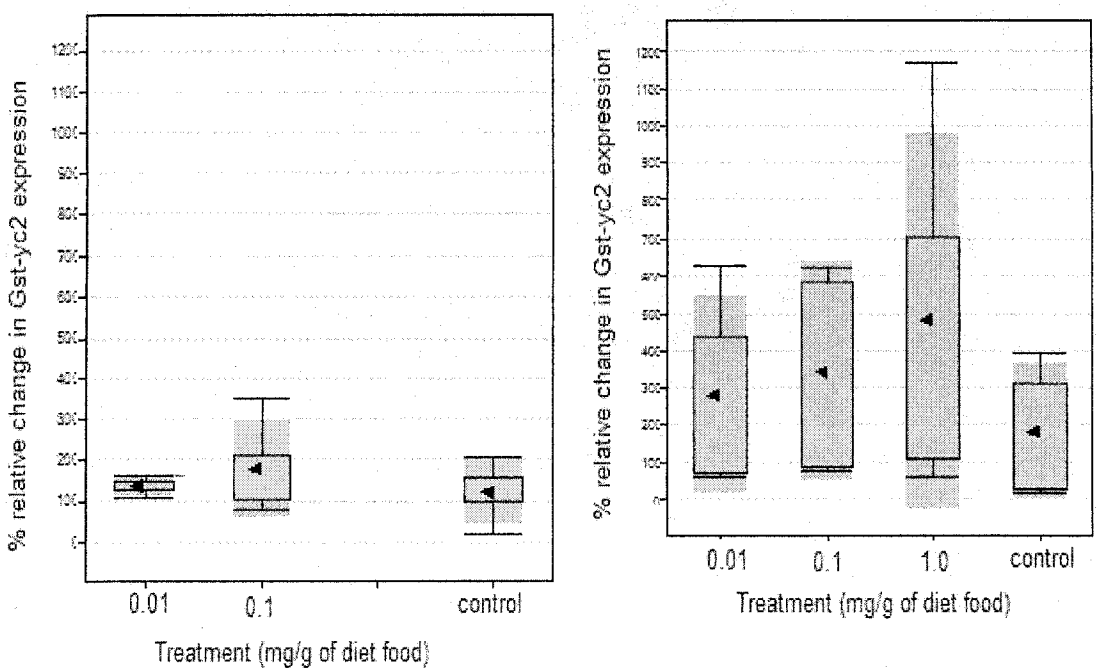
Effects of BaP on Gst-yc2 Expression

Significant changes ($p < 0.05$) in gene expression of *Gst-yc2* in rat livers after 12-weeks (Figure 25b) of exposure to control, 0.01, 0.1, 1.0 mg/g of BaP in the diet. The increase in expression of *Gst-yc2* observed after 12-weeks of exposure was dose dependent. Highest expression was observed for 1.0 mg/g of BaP in diet (Figure 25b). The low, medium and high dose showed approximately two, three and four old increases in expression over the control average.

On the other hand after two weeks of exposure, only a slight increase in expression of *Gst-yc2* over the control was observed; 37 % and 78 % increase for the low and high dose, respectively.

The observed increased was not significant at $p < 0.05$. No significant difference in *Gst-yc2* expression was observed after 2-weeks (Figure 25a) of exposure to control, 0.01, and 0.1 mg/g of BaP in the diet. A high variation in expression of *Gst-yc2* was observed within treatment and control groups in 2- and 12-weeks weeks exposed rats as indicated by the box plot.

The results clearly show differences in the temporal expression of *Gst-yc2* between the 2- and 12-weeks time period. Higher expression of *Gst-yc2* at the 12-weeks time period is consistent with the induction of such genes, whose protein products are responsible for detoxification and metabolism of oxidants.



(a) (b)

Figure 25. Relative Change in Gene Expression of *Gst-yc2*. (a) Percent change in *Gst-yc2* after 2-weeks Exposure to 0.01 and 0.1 mg of BaP/g of Diet. (b) Percent Change in *Gst-yc2* after 12-weeks Exposure to 0.01, 0.1 and 1.0 mg of BaP/g of Diet

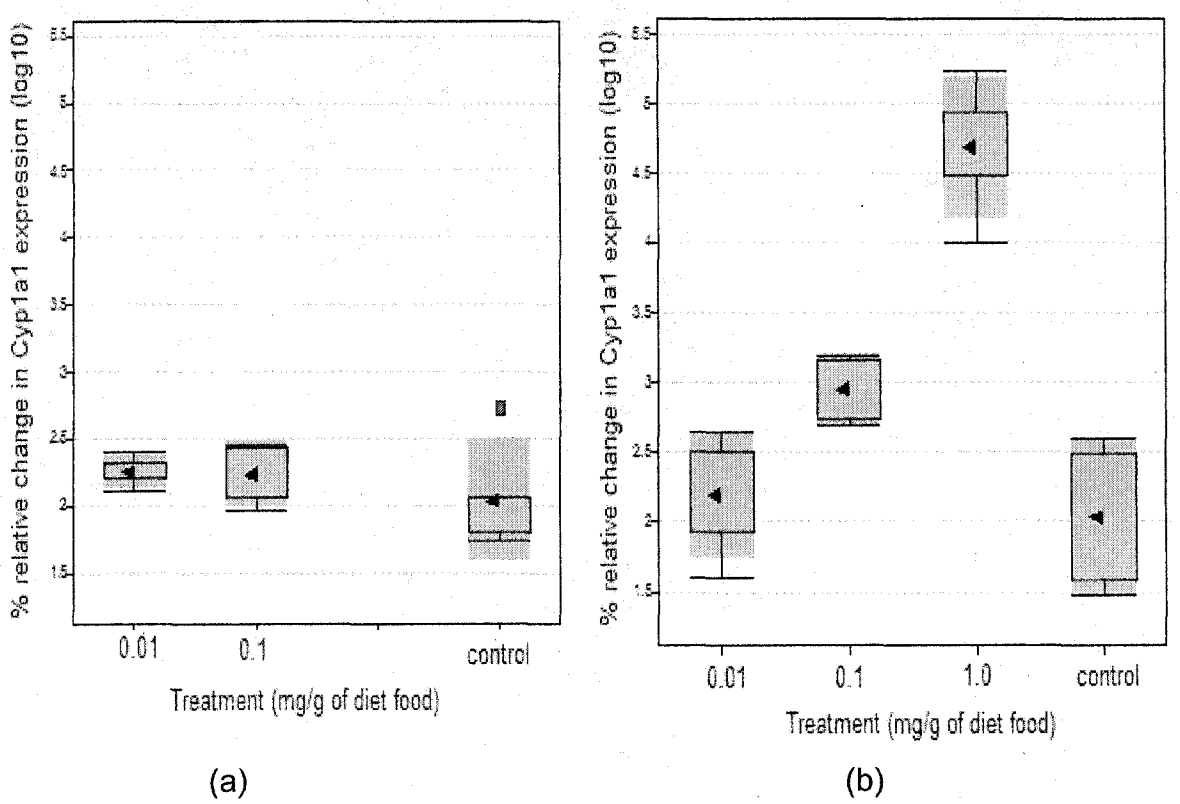
Effect of BaP on Cyp1a1 and Cyp1a2 Expression

We observed significant differences over control animals ($p < 0.05$) in expression of *Cyp1a1* and *Cyp1a2* in rat livers after 12-weeks (Figure 26b & 27b) of exposure to 0.1 and 1.0 mg/g of BaP in the diet. Bonferroni post-hoc analysis also reveals significant differences ($p < 0.05$) between controls of *Cyp1a1* and *Cyp1a2* expression compared to the high dose (1.0 mg/g-diet). For *Cyp1a1* expression, the low, medium, and high dose showed 2, 10, and 74 fold increases over the control expression, while *Cyp1a2* showed 2, 3, and

9 fold increases in expression over the control values, respectively. No significant difference ($p < 0.05$) in expression of *Cyp1a1* and *Cyp1a2* was observed between treatment groups after 2-weeks (Figure 26a and 27a) of exposure to BaP. Only slight increases in the expression of *Cyp1a1* (79 % and 88% for low and high dose) was observed. For *Cyp1a2,* the expression after 2-weeks of exposure was less than 2-fold in both low and high dose. Furthermore, the expression of *Cyp1a1* and *Cyp1a2* was consistently higher than controls in all animals exposed to the medium dose (0.1 mg/g) and high dose (1mg /g) after 12-weeks of exposure to BaP. This shows that in addition to temporal regulation, the amount and nature of the dose as discussed earlier may be important for the expression of detoxification enzymes such as *Cyp450 1a1* and *1a2*.

Effect of BaP on p53 and Bcl-2 Expression

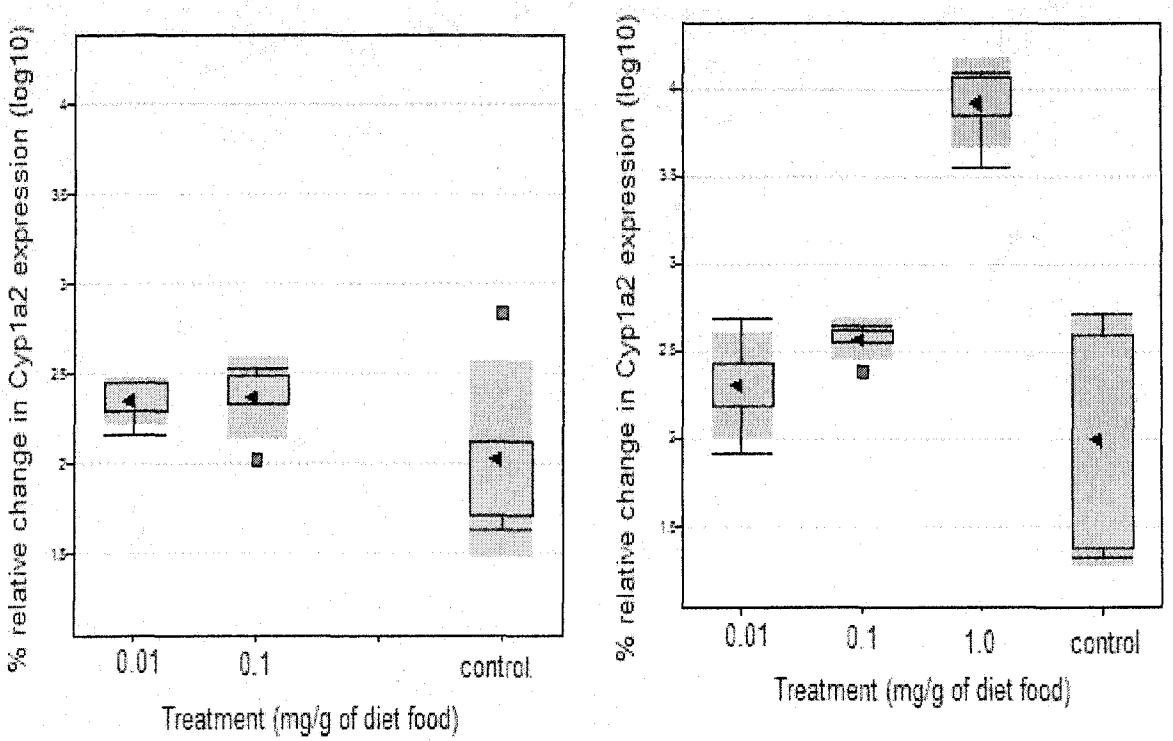
Significant ($p < 0.05$) changes in *p53* expression were observed in rat livers after 12-weeks (Figure 28b) of exposure to control, 0.01, 0.1, and 1.0 mg/g of BaP in the diet. After 12-weeks, the highest increase (3-fold) in expression over controls was observed in the low dose (0.01 mg/g), followed by a 2-fold increase in both the medium (0.1 mg/g) and high (1.0 mg/g) dose, respectively. Similarly, after 2-weeks, significant differences in *p53* expression were observed in rat livers exposed to 0.1 mg/g of BaP in their diet (Figure 28a).



(a) (b)

Figure 26. Relative Change in Gene Expression of *Cyp1a1*. (a) Percent Change in *Cyp1a1* after 2-weeks Exposure to 0.01 and 0.1 mg of BaP/g of Diet. (b) Percent Change in *Cyp1a1* after 12-weeks Exposure to 0.01, 0.1 and 1.0 mg of BaP/g of Diet

A threefold increase in *p53* expression was observed after exposure to the BaP dose of 0.1 mg/g of diet food after 2-weeks. Only a slight mean increase in *p53* expression was observed in the 0.01 mg/g of BaP treatment group. Similar changes in expression were observed for *Bcl-2*, which showed significant changes ($p < 0.05$) in gene expression in all 4 doses (control, 0.01, 0.1, and 1.0 mg/g-diet) after 12-weeks (Figure 29b) of exposure to BaP. Bonferroni post-hoc analysis reveals significant differences ($p<0.05$) from control for the medium and high dose after 12-weeks of exposure.

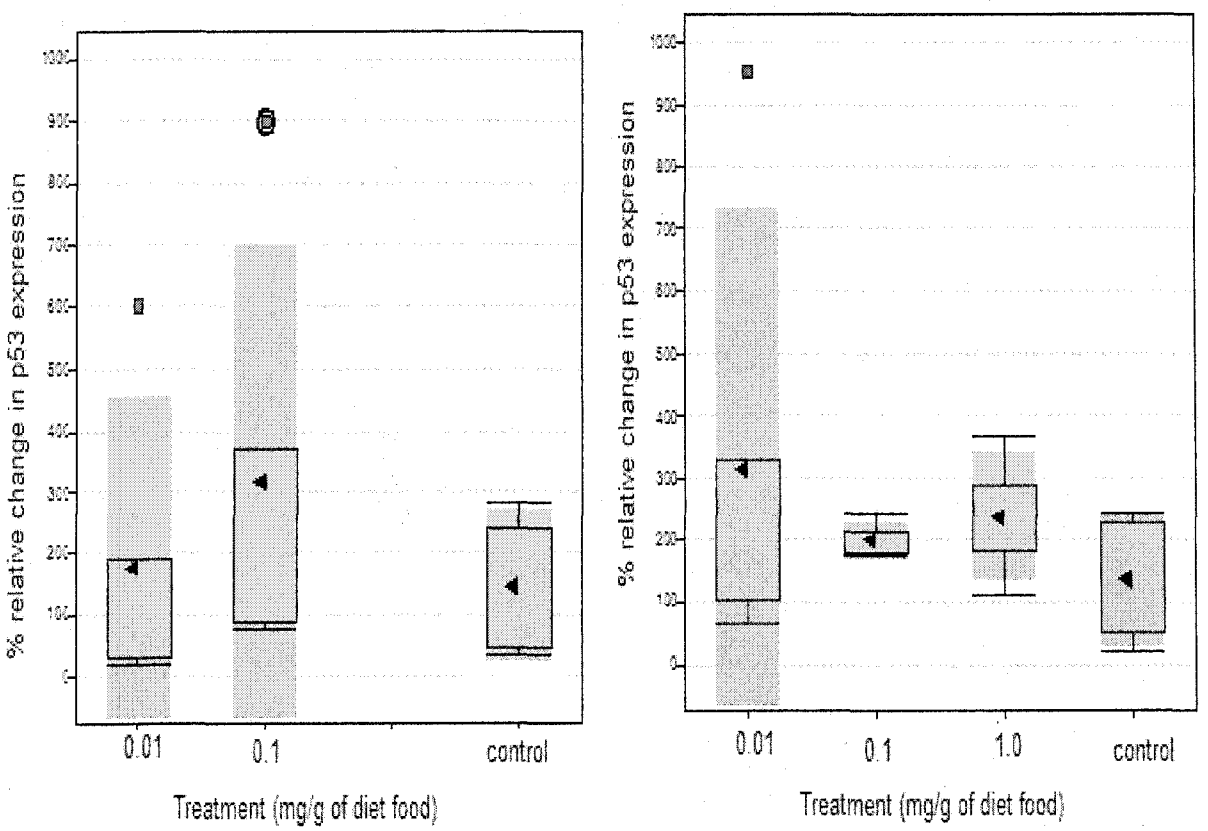


(a) (b)

Figure 27. Relative Change in Gene Expression of *Cyp1a2*.. (a) Percent Change in *Cyp1a2* after 2-weeks Exposure to 0.01 and 0.1 mg of BaP/g of Diet. (b) Percent Change in *Cyp1a2* after 12-weeks Exposure to 0.01, 0.1 and 1.0 mg of BaP/g of Diet

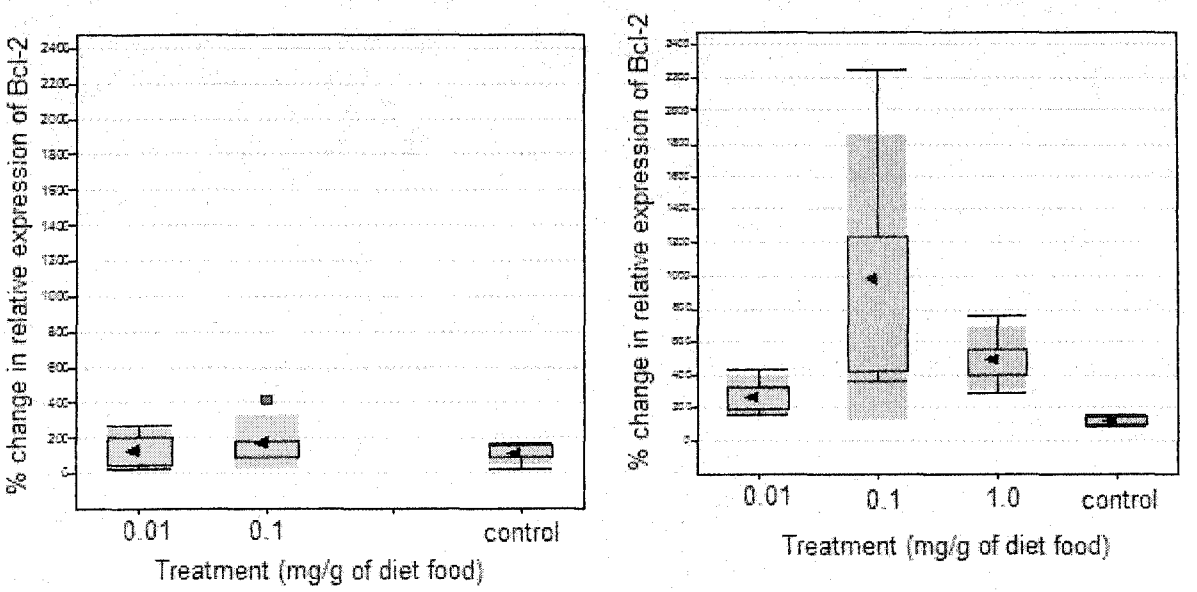
On the other hand, *Bcl-2* expression increased by 2, 9.5, and 4.5 fold for the low, medium and high dose, respectively after 12-weeks of exposure. A slight increase in mean expression of *Bcl-2* was observed during the 2-weeks exposure period to BaP, however, this differences was not significant ($p < 0.05$) (Figure 29b). These results are in agreement with previous studies that show BaP induces apoptotic genes such as *p53*. A study by Pei et al. (1999) demonstrated that BaP activates the human *p53* gene through induction of the nuclear factor kB activity. Exposure of A549 and NIH 3T3

cells to BaP resulted in an increase in *p53* mRNA levels. A related study by Salas and Burchiel (1998) also found that BaP induces apoptosis in Daudi cells via a mechanism that involves *Bcl-2* family proteins, potentially secondary to the formation of p450-derived metabolites and alterations.



(a) (b)

Figure 28. Relative Change in Gene Expression of *p53.* (a) Percent Change in *p53* after 2-weeks Exposure to 0.01 and 0.1 mg of BaP/g of Diet. (b) Percent Change in *p53* after 12-weeks Exposure to 0.01, 0.1 and 1.0 mg of BaP/g of Diet



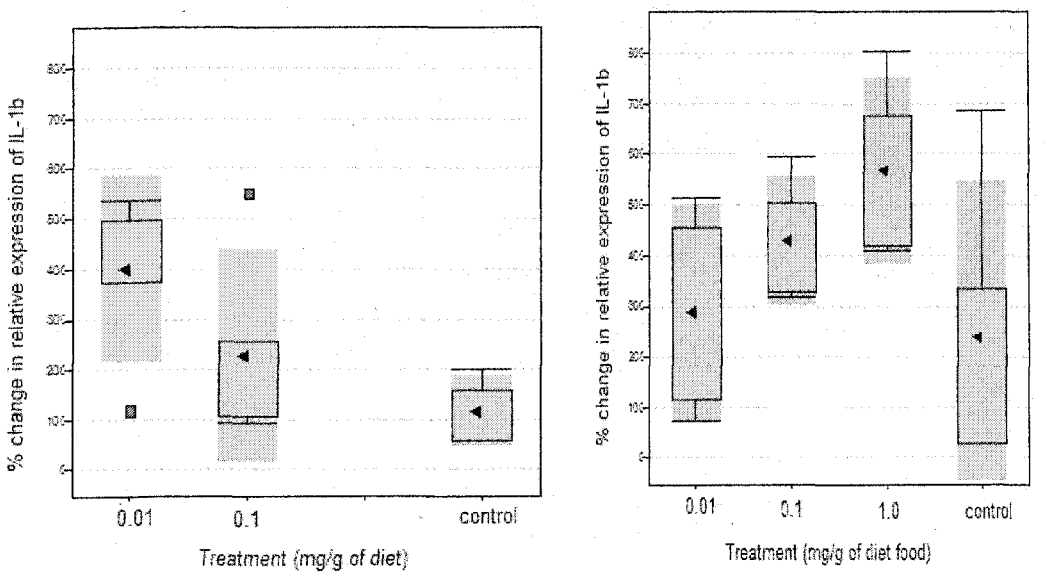
(a) (b)

Figure 29. Relative Change in Gene Expression of *Bcl-2*. (a) Percent Change in *Bcl-2* after 2-weeks Exposure to 0.01 and 0.1 mg of BaP/g of Diet. (b) Percent Change in *Bcl-2* after 12-weeks Exposure to 0.01, 0.1 and 1.0 mg of BaP/g of Diet

Effect of BaP on IL-1β Expression

Significant differences ($p < 0.05$) in expression of *IL-1β* were observed in rat liver mRNAs after 12-weeks (Figure 30b) of exposure to BaP. The increase in expression after 12-weeks was also found to be dose-dependent. For example, the low, medium and high dose after 12-weeks showed approximately 2.5, 4, and 5.5 fold increases in expression over the control value. In addition, significant differences were also observed in the 2-weeks exposed (Figure 30a) animals; both the low and high dose levels of BaP caused 4 and 2-fold induction of *IL-1b* over the controls. For the 2-weeks treatment groups, a wide range of variations in the mean values was

observed, with one or two outliers as shown by the red dots in Figure 30a. These results are in line with the role of *IL-1b* as ubiquitous molecule that plays a role in the inflammatory response and the immune system. Expression of *IL-1b* is linked to increases in Intracellular calcium levels (Davila *et al*, 1995), and since BaP and other PAHs can cause increase in intracellular calcium levels, they may also cause increase in the *IL-1b* gene expression.



(a) (b)

Figure 30. Relative Change in Gene Expression of *IL-1β*. (a) Percent Change in *IL-1β* after 2-weeks Exposure to 0.01 and 0.1 mg of BaP/g of Diet. (b) Percent Change in *IL-1β* after 12-weeks Exposure to 0.01, 0.1 and 1.0 mg of BaP/g of Diet.

In Vivo Rat Feeding Experiments

In vivo rat feeding study was conducted as discussed in the materials and methods section. BaP was administered to rats through the diet for 2-and 12-

weeks time period. A HPLC assay was run on each feed sample prior to administration to ensure the homogeneity and potency of BaP in feed before it was used. A standard plot was constructed based on the peak area to achieve the more accurate result. The linear equation of y = 23953x and R^2 = 0.9996 were obtained with accuracy from 99% to 104%. The standard deviation of three injections for all levels of standards was less than 0.5 % and the highest percent relative standard deviation (%RSD) for all levels of standard injections was less than 0.5 %. This data indicates the method of determining BaP concentrations in the diets is accurate, precise and suitable for the analysis. Representative chromatograms of feed samples are shown in Figure 31.

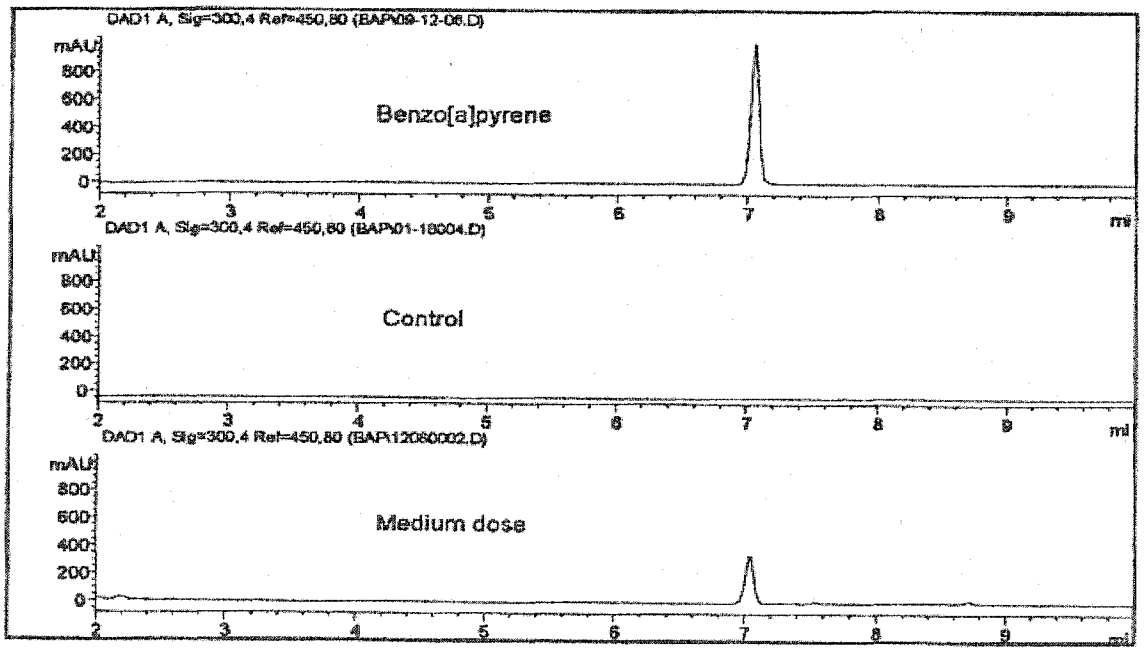


Figure 31. LC Chromatograms of Feed Samples.

During the *in vivo* rat feeding study period, the body weight of the rats and amount of food consumed was monitored constantly. No significant change in the body weight of the rats was observed after two weeks of exposure to 0.01 and 0.1 mg/g of BaP. However, after 12-weeks of exposure to BaP, the average weight gain of the high dose (1.0 mg/g) group was significantly lower than other groups throughout the 12-weeks study period. The Animals in high dose group appeared slower and abnormal fur loss was observed. They were less active than the other groups and internal organs showed some morphological differences from the control animals.

The *in vivo* rat experiment average weight gain results are presented in Table 12. The average weight gain results normalized to feed consumption are presented in Figures 32 and 33. Similar changes in body weight with dose has been observed in other studies (Chiang, 2001, Ramesh et al, 1999)

At the end of each *in vivo* rat feeding study period (2-and 12-weeks), animals were sacrificed, blood collected through heart puncture and various organs (liver, lungs, kidneys, spleen, thymus, and heart) were collected. The organs were weighed immediately after extraction to obtain the fresh weight of each organ. Significant differences in weight of the organs between control and treated animals were observed for each time point.

The average weights of the various organs compared to the control animals are in Figures 34 through 39.

Table 12

In Vivo Dosage and Average Weight Gain Results

	2-week Exposure			12-week Exposure												
Dosing Period (Weeks)	1	2	Average Weight Gain (gm)	1	2	3	4	5	6	7	8	9	10	11	12	Average Weight Gain (gm)
High Dose (1.0mg)	NA	NA	NA	10.1	12.5	3.8	2.0	3.6	3.5	8.0	9.5	10.0	9.5	8.6	8.5	7.5±3.3
Medium Dose (0.1mg)	23.2	20.1	21.7±2.2	20.1	13.7	12.8	11.8	9.0	9.5	9.1	8.5	7.8	10.5	9.6	10.5	11.1±3.3
Low Dose (0.01mg)	24.6	21.7	23.2±2.3	22.5	21.3	9.5	15.1	11.3	10.1	9.5	8.1	6.3	7.5	8.1	9.4	11.6±5.3
Control	19.9	19.3	19.6±0.4	20.9	20.1	15.2	14.7	10.5	11.3	9.9	10.1	8.8	8.3	8.9	9.5	12.4±4.4

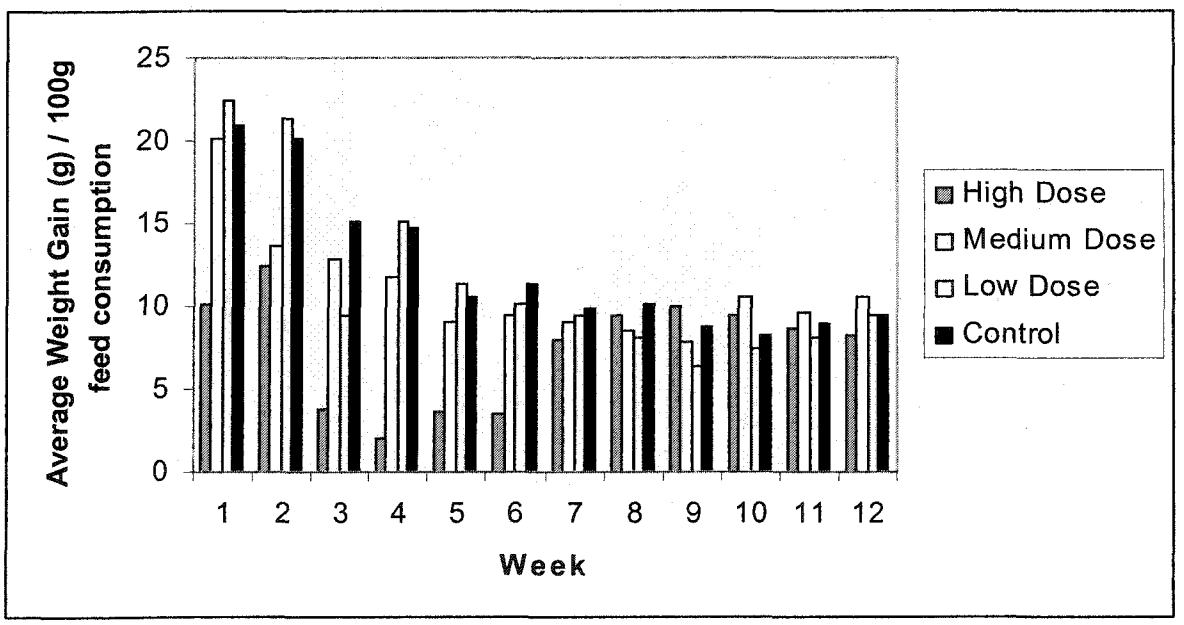


Figure 32. *In Vivo* Dosage and Average Weight Gain Normalized for Feed Consumption

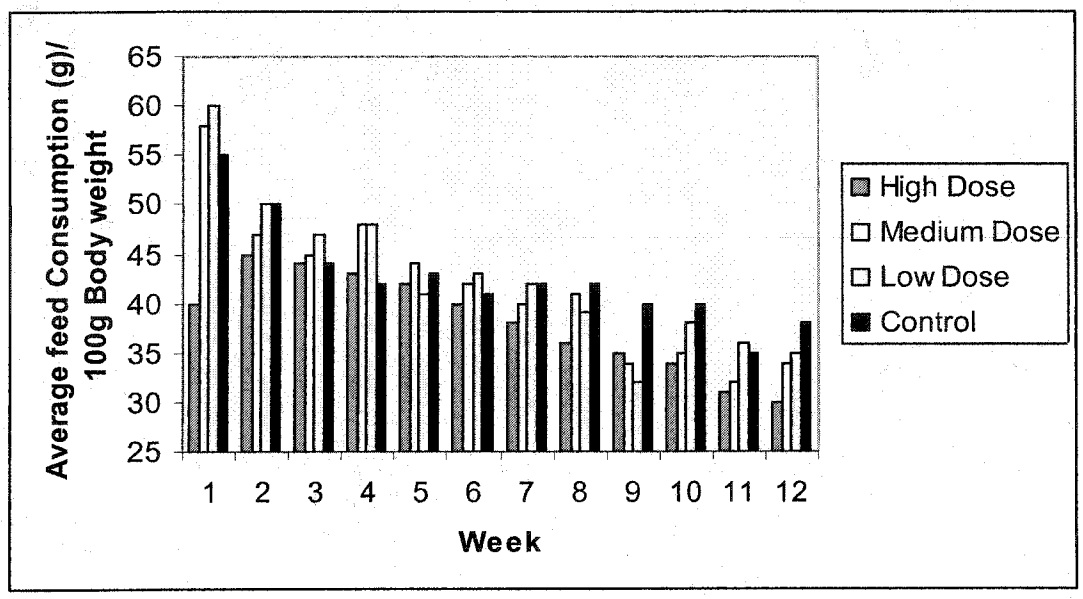
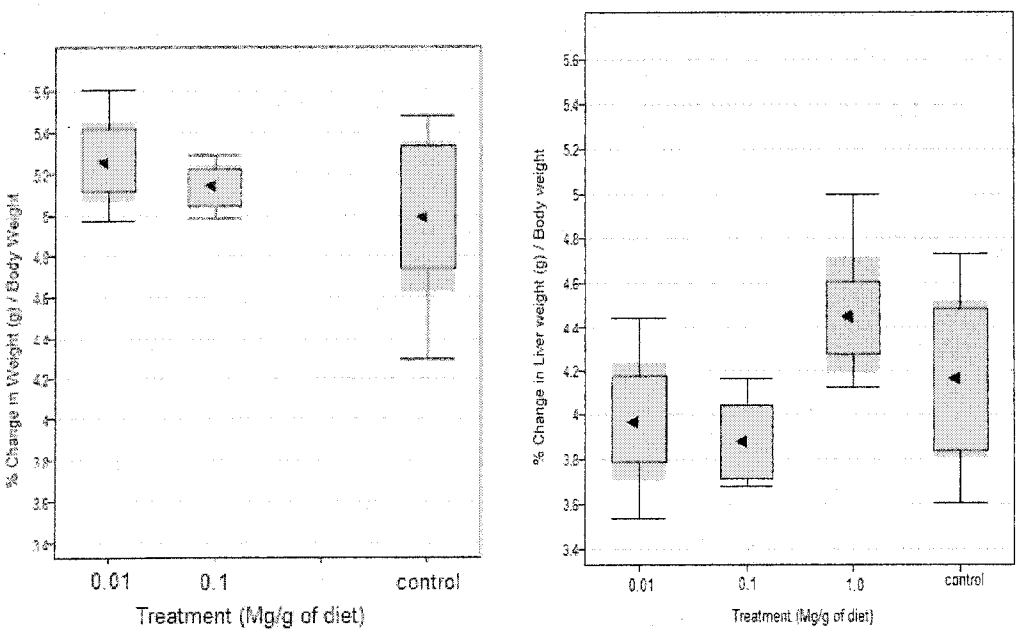


Figure 33. *In Vivo* Dosage and Average Feed Consumption Normalized for Body Weight



(a) (b)

Figure 34. Percent Change in Liver Weight. (a) 2-weeks of Exposure to BaP Control, 0.01, 0.1, and 1.0 BaP mg/g of diet. (b) 12-weeks of Exposure to Control, 0.01, 0.1, and 1.0 BaP mg/g of diet.

As shown in Figure 34, changes in liver size were observed for both the 2-and 12-weeks treatment period. A mean increase of 3-5 % in liver weight was observed in liver weight after 2-weeks of exposure (Figure 34a) to low (0.01 mg/g) and high dose (0.1 mg/g) of BaP. On the other hand, after 12-weeks (Figure 34b) of exposure, a 5-8 % decrease in the size of the liver was observed in the low (0.01 mg/g) and Medium (0.1 mg/g) dose of BaP, whereas a 6% increase in the size of the liver was observed in high dose treatment groups (1mg/g). The liver is an important organ in protection from oxidative damage. It plays a major role in the breakdown of potentially toxic lipophilic toxins through the action of oxidation and reduction enzymes such as cytochrome p-450 enzymes.

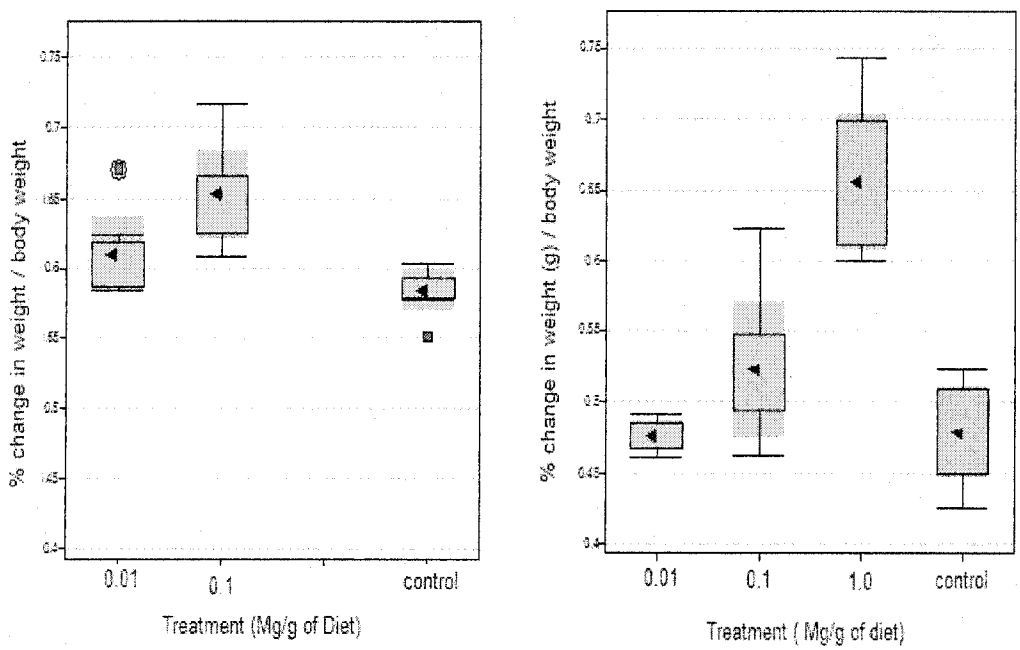


Figure 35. Percent Change in Lung Weight. (a) 2-weeks of Exposure to BaP Control, 0.01, 0.1, and 1.0 BaP mg/g of diet. (b) 12-weeks of Exposure to Control, 0.01, 0.1, and 1.0 BaP mg/g of diet.

For the lungs data, as shown in Figure 35a, a dose-dependent increase in size of the lungs was observed; 4 % in the low dose (0.01mg/g) and 11.0 % in the high dose (0.1mg/g). For the 12-weeks exposure (Figure 35b), lung size also showed no change from the control animals in the low dose (0.01mg/g) treated groups, whereas, in the medium (0.1 mg/g) and high (1.0 mg/g) dose, the increase in lung size was dose-dependent; 9% and 37% increase in lung size in the medium and high dose after 12-weeks of exposure. The induction of Cyp1a1 and Cyp1a2 in human lungs by BaP is associated with lung cancer (Wei et al, 2001).

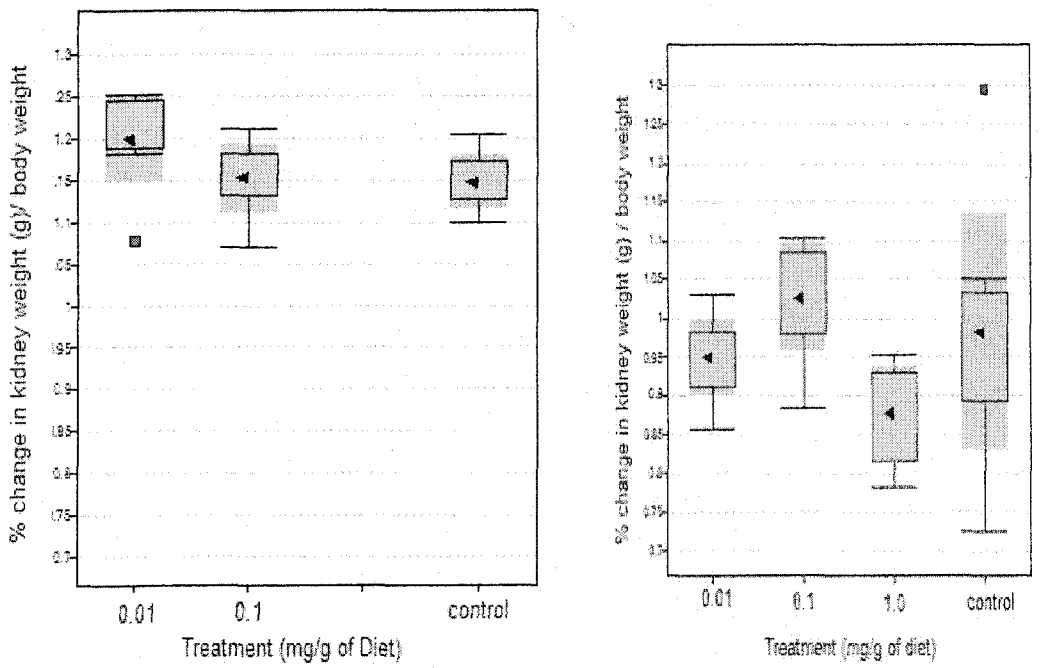


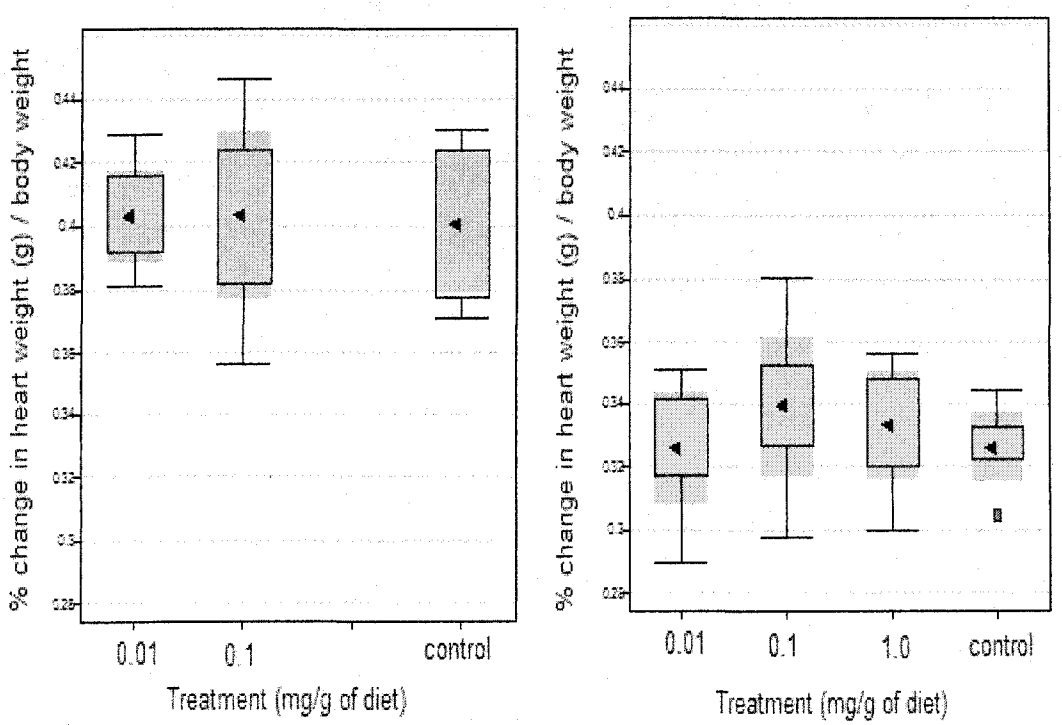
Figure 36. Percent Change in Kidney Weight. (a) 2-weeks of Exposure to BaP Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet. (b) 12-weeks of Exposure to Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet.

As an excretory organ, the kidney plays an important role in detoxification and metabolism of xenobiotic compounds and drugs. An increase of about 4 % in the kidney weight was observed for the low and high dose after 2-weeks (Figure 36a). On the other hand, after 12-weeks, the kidney size decreased in the high (1.0 mg/g) and low (0.01 mg/g) dose of animals by 3-8%, and only increased by 4% in the medium (0.1 mg/g) of BaP treatment (Figure 36b). In addition, no significant increase in the size of the heart was observed after 2-weeks of exposure to low and high dose of BaP (Figure 37a). For the 12-weeks exposure period, the mean size of the heart for the low dose and the control remain the same, whereas a 4% increase in heart size was observed for the medium (0.1 mg/g) and high (1.0 mg/g) dose of BaP after 12-weeks of exposure (Figure 37b).

BaP has been implicated in cardiovascular diseases such as atherosclerosis (Yan et al, 2000) in animals ranging from chickens to rats (Ou and Ramos, 1992). Also, Yan et al. (1998) has detected BaP-related adducts in atherosclerotic arteries. COX-2 gene is expressed in atherosclerotic lesions and BaP has been shown to induce the COX-2 gene, which in epithelial cells inhibits apoptosis, and increases the invasiness of malignant cells, favoring tumor formation. Prolong period of dosing may be needed to observe these effects in the heart.

A change in mean spleen weight was observed after 2-weeks of exposure to BaP (Figure 38a). A moderate mean increase in weight of 2-3% was observed for the low dose group, while there was a 12% increase in the mean

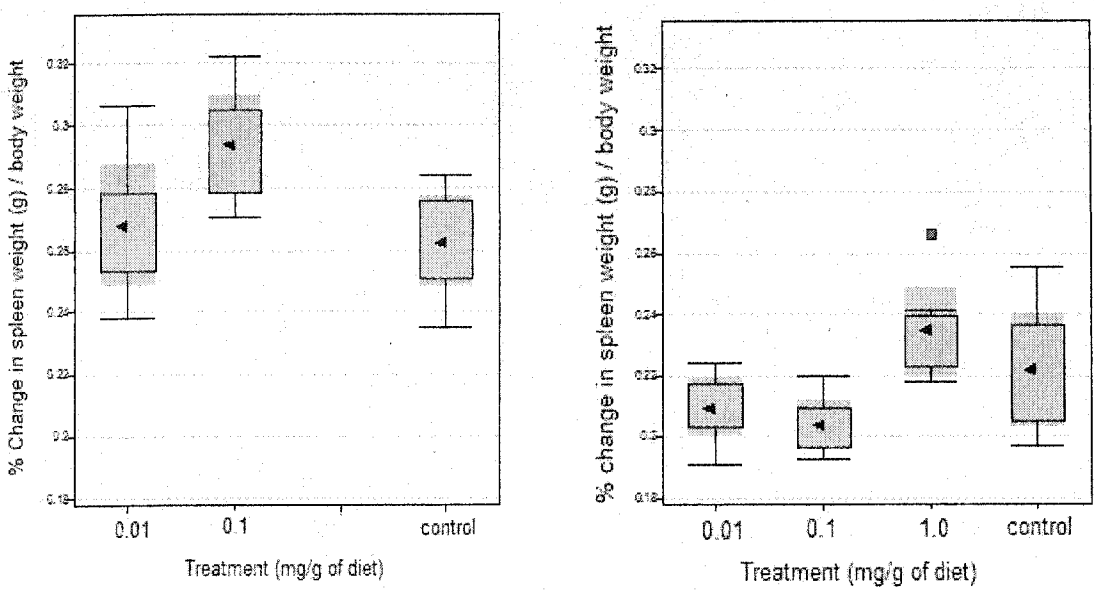
weight of the spleens in high dose group after 2-weeks of exposure to BaP. However, after 12-weeks of exposure (Figure 38b), a 6-8% decrease in mean weight was observed for the low (0.01 mg/g) and medium (0.1 mg/g) dose of BaP. High dose (1 mg/g) of BaP, on the other hand, showed a 6% mean increase in spleen weight after 12-weeks of exposure (Figure 38b).



(a) (b)

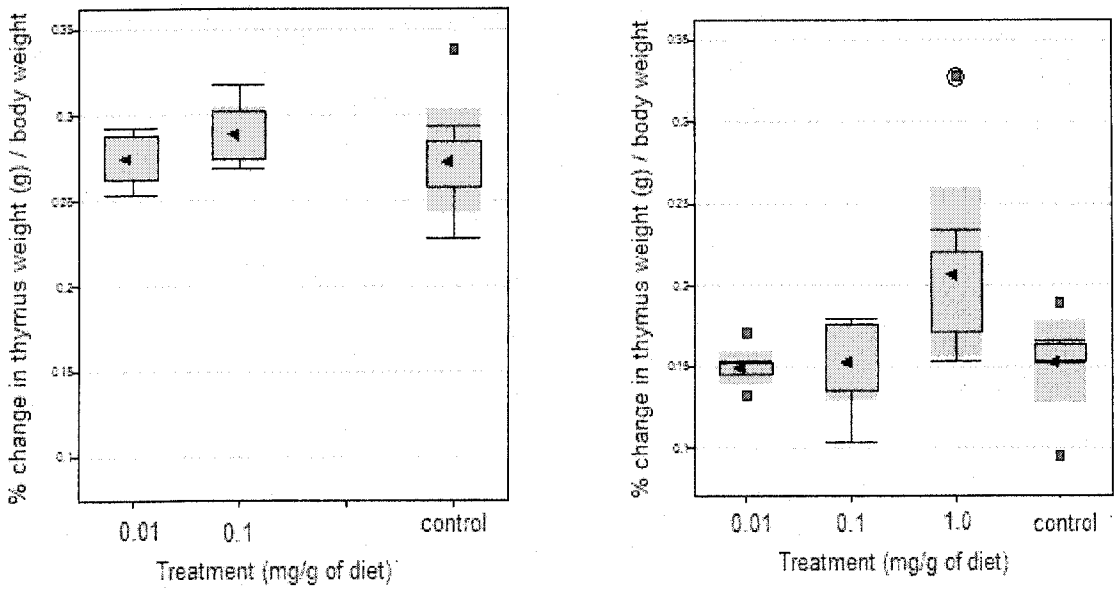
Figure 37. Percent Change in Heart Weight. (a) 2-weeks of Exposure to BaP Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet. (b) 12-weeks of Exposure to Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet.

In the case of the thymus, only the high dose (0.1 mg/g) of BaP showed a mean increase of 6% over the control animals after 2-weeks of exposure (Figure 39a).



(a) (b)

Figure 38. Percent Change in Spleen Weight. (a) 2-weeks of Exposure to BaP Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet. (b) 12-weeks of Exposure to Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet.



(a) (b)

Figure 39. Percent Change in Thymus Weight. (a) 2-weeks of Exposure to BaP Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet. (b) 12-weeks of Exposure to Control, 0.01, 0.1, and 1.0 BaP mg/g of Diet.

After 12-weeks of exposure, there was no difference between the mean weights of the control thymus and the thymus from the low (0.01 mg/g) and medium (0.1 mg/g) dose groups. Conversely, there was 36% increase in the thymus weight of rats exposed to the high dose of 1 mg/g of BaP for 12-weeks (Figure 39b)

CHAPTER V

CONCLUSION

Gene expression profiling using DNA microarrays and validation by Real-time Quantitative PCR as shown in this study may be instrumental in defining molecular and biochemical mechanisms of the action of toxicants. This study was focused on the analysis of BaP induced gene expression changes in Fisher 344 rat livers utilizing these expression profiling techniques. The ultimate goal of our research project was to identify and characterize biomarkers or signature genes associated with BaP toxicity. This may also be useful in understanding the relationship between environmental exposures and human disease susceptibility. The result of the BaP microarray analysis experiment has provided answers to the following questions: 1) reliability of microarray GeneChips as a tool for risk assessment and gene expression profiling, 2) the nature of the dose-response and temporal regulation of gene expression by BaP, 3) the gene expression profile of markers associated with BaP exposure, and 4) functional understanding of the biological complexity associated with mechanism of toxicity of BaP and other related compounds.

We show in our experimental data that microarrays can act as a reliable tool for toxicogenomics studies. Chip to chip comparisons of technical replicates showed 96 % similarity in the genes expressed. The only variation that was observed in the data comes from normal biological variation between individual animals. Furthermore, the microarray experiments also showed that

BaP causes differential expression of a large family of genes in the rat livers under acute as well as sub-chronic conditions. Differential expression was observed for all doses and time points used. In addition, the results show that both dose and exposure time intervals are important factors in determining not only the number of genes that are regulated, but also the extent to which these genes are increased or decreased. A strong variation in the expression profile of genes belonging to different functional categories was observed from the data. A good number of the genes were induced early, while others showed late induction. A temporal separation in the expression of several genes was observed from the data. Furthermore, at the low dose exposure level, an adaptive response was observed for a good number of genes, indicating some form of tolerance or a return to normal or sub-normal expression levels.

Additionally, the microarray data revealed changes in gene expression for several interesting physiological pathways including the p450 detoxification, redox, cell signaling, metabolism, oncogenes, apoptosis and steroid biosynthesis pathways. A wide range of genes and express sequence tags (ESTs) previously unknown to be induced by BaP were uncovered. Other genes that are known to be regulated by BaP were confirmed by the data. The data also reveals other genes or ESTs that may be involved in the regulation of these genes. Through construction of pathway maps, the relationships between the exposure to BaP and the molecular mechanisms for disease susceptibility were elucidated. Disruption of a good number of

genes in these pathways is implicated in biochemical processes leading to cancer and other health effects in humans. Also, we have used a real-time quantitative Taqman PCR to validate and quantify the changes in expression of 6 target genes of interest, that were previously identified from Affymetrix microarray rat RG-U34A gene chips containing 8,800 genes. A significant change in gene expression over control groups ($p < 0.05$) was observed after 2-and 12-weeks of exposure to BaP for the genes of interest.

Gene expression profiling using DNA microarrays represent a snapshot of the gene transcription occurring at a particular time and within a particular tissue. Toxicity, on the other hand, represents a continuum of possible effects governed by both temporal and spatial factors that depend on the exposure conditions. The environmental conditions that define exposure to a contaminant may also cause significant variability in gene expression data. Therefore, future studies that take into account a wide range of time points, tissue or cell types, dosage routes and doses are required. Dose-response relationships need to be characterized in order to make sense of gene expression patterns derived from microarrays. The use of different dosing route would help determine better routes of bioavailability. The use of different tissues or cells would allow comparison of the spatial differences in gene expression patterns, and therefore give a clearer picture of the transcription profile of specific genes in different tissues. Also, an increase number of time points allow capture of the temporal phase distinction and regulation of genes or gene families in response to specific toxicants or classes of toxicants. The

use of better or improved mathematical and statistical techniques specifically designed for gene expression analysis would enhance interpretation of the data. One such promising technique involves analysis of gene expression data using functional principal components (Barra, 2003). This method considers gene profiles as continuous curves, and applies a mathematical function analysis to gather genes with similar variations and extract characteristic parameters of gene profiles. Finally, proteomic techniques such as western blot, mass spectrometry, proteomic arrays, as well as RNA interference should be utilized in future studies to determine protein profiles and mechanisms of action of potential genes.

Integration of these tools will help in elucidating the various cellular networks in biological systems, and help in predicting the toxicological outcome of contaminant exposure in model and non-model organisms. This may eventually create a good knowledge base for protecting human health.

APPENDIX

A1

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 1

Probe Set ID	Gene Name	Molecular Function Description
AB002558_at	glycerol-3-phosphate dehydrogenase 1 (soluble)	glycerol-3-phosphate dehydrogenase (NAD+) activity
AF025308_	RT1 class I, A3, RT1 class Ib, locus	
D13667cds_s_	alanine-glyoxylate aminotransferase	serine-pyruvate transaminase activity alanine-glyoxylate transaminase activity
D14989_f_at	MRNA for hydroxysteroid sulfotransferase subunit, complete cds	
E03358cds_at	proteasome (prosome, macropain) subunit, alpha type 2	endopeptidase activity, peptidase activity hydrolase activity
E12286cds_at	GM2 ganglioside activator protein	
M14776_f_at	Cytochrome P450, subfamily IIC6	monooxygenase activity, oxidoreductase
M21476_s_at	prolyl 4-hydroxylase, beta polypeptide	protein disulfide isomerase activity procollagen-proline 4-dioxygenase activity electron transporter activity
rc_AA92547	cell division cycle 42 homolog (S. cerevisiae)	GTPase activity
rc_AA99648	aldehyde dehydrogenase family 3, subfamily A2	aldehyde dehydrogenase activity aldehyde dehydrogenase (NAD) activity oxidoreductase activity
rc_AI171734_s_at	fumarate hydratase 1	catalytic activity, fumarate hydratase activity, lyase activity
S35751_f_at	3-alpha-hydroxysteroid dehydrogenase	
S81448_s_at	steroid 5 alpha-reductase 1	3-oxo-5-alpha-steroid 4-dehydrogenase activity, oxidoreductase activity
X53054_g_at	RT1 class II, locus Bb RT1 class II, locus Db1	
X70223_at	peroxisomal membrane protein 2	peroxisome organization and biogenesis
X72792cds	alcohol dehydrogenase 1	alcohol dehydrogenase activity, zinc-dependent,zinc ion binding, oxidoreductase

A2

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 2

Probe Set ID	Gene Name	Molecular Function Description
AA685903_at	Similar to Endoplasmin precursor (Endoplasmic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorphic tumor rejection antigen 1) (Tumor rejection antigen gp96) (LOC362862), mRNA	
AB012933_g_at	acyl-CoA synthetase long-chain family member 5	magnesium ion binding,catalytic activity long-chain-fatty-acid-CoA ligase activity
AF023302_s_at	flotillin 2	
AFFX_Rat_GAPDH_5_at	glyceraldehyde-3-phosphate dehydrogenase	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity oxidoreductase activity
AJ223355_at	solute carrier family 25 (mitochondrial carrier; dicarboxylate transporter), member 10	
D00913_g_at	intercellular adhesion molecule 1	defense response, cell adhesion, cell-cell adhesion
D13921_s_at	acetyl-coenzyme A acetyltransferase 1	acetyl-CoA C-acetyltransferase activity acyltransferase ,transferase activity
D13978_s_at	argininosuccinate lyase	catalytic activity, argininosuccinate lyase activity, lyase activity, urea cycle arginine biosynthesis
D37934_g_at	Similar to proteasome 26S non-ATPase subunit 2 (LOC287984), mRNA	
D85435_g_at	protein kinase C, delta binding protein	protein kinase C binding, regulation of cell growth
D87839_g_at	4-aminobutyrate aminotransferase	4-aminobutyrate transaminase activity transaminase activity, transferase activity
D87991_at	galactose transporter	
D90401_g_at	afadin	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 2 – Continued

E01050cds_g_at	alanine-glyoxylate aminotransferase	serine-pyruvate transaminase activity alanine-glyoxylate transaminase activity transaminase activity, transferase activity
E03859cds_s_at	RAB11a, member RAS oncogene family	GTP binding, protein transporter activity GTPase mediated signal transduction
E13557cds_s_at	cysteine-sulfinate decarboxylase	sulfinoalanine decarboxylase activity lyase activity, carboxy-lyase activity, amino acid metabolism, taurine biosynthesis
J00735_at	fibrinogen, gamma polypeptide	inflammatory response, blood coagulation
J05210_at	ATP citrate lyase	magnesium ion binding, catalytic activity ATP citrate synthase activity, citrate (Si)-synthase activity, ATP binding, transferase activity, lyase activity, ligase
K00750exon#2-3_at	Similar to Cytochrome c, somatic (LOC365738), mRNA	
M15481_g_at	insulin-like growth factor 1	hormone activity, growth factor activity
M17412_at	growth and transformation-dependent protein	
M26247_at	coagulation factor 9	coagulation factor IXa activity serine-type endopeptidase activity chymotrypsin activity, trypsin activity calcium ion binding, hydrolase activity
M62832_at	plasminogen	
M64755_at	cysteine-sulfinate decarboxylase	sulfinoalanine decarboxylase activity lyase activity, carboxy-lyase activity
M64795_f_at	RT1 class Ib, locus Aw2	
X13058_at	Tumor suppressor protein p53	DNA binding, transcription factor
M86870_at	protein disulfide isomerase related protein (calcium-binding protein, intestinal-related)	protein disulfide isomerase activity
M95768_at	chitobiase, di-N-acetyl-	catalytic activity, hydrolase activity,

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 2 – Continued

rc_AA799279	Similar to mitochondrial carrier homolog 2 (LOC295922), mRNA	
rc_AA7994	Hypothetical LOC293114 (LOC293114), mRNA	
rc_AA799571	Similar to RIKEN cDNA 1110001M20 (LOC298308), mRNA	
rc_AA817854_s_at	ceruloplasmin	ferroxidase activity, copper ion binding oxidoreductase activity
rc_AA858573	spp-24 precursor	
rc_AA858640	heat shock protein 60 (liver)	
rc_AA892799	Similar to glyoxylate reductase/hydroxypyruvate reductase (LOC298085), mRNA	
rc_AA892861_at	Similar to RIKEN cDNA 2610528J11 (LOC362576), mRNA	
rc_AA893905_g_at	tripartite motif protein 39	
rc_AI011998_at	DnaJ (Hsp40) homolog, subfamily B, member 9	chaperone activity
rc_AI013194	eukaryotic translation initiation factor 5	translation initiation factor activity GTP binding
rc_AI044341_	valosin-containing protein	nucleotide binding, ATP binding hydrolase activity
rc_AI169735	cytochrome P450IIB3	monooxygenase activity, oxidoreductase
rc_AI170568_s_at	dodecenoyl-coenzyme A delta isomerase	catalytic activity, dodecenoyl-CoA delta-isomerase activity
rc_AI235890_	MHC class Ib RT1.S3	
rc_AI237007	electron-transferring-flavoprotein dehydrogenase	
S45663_at	glycoprotein, synaptic 2	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 2 – Continued

S75019_at	Similar to aldehyde dehydrogenase family 7, member A1; aldehyde dehydrogenase 7 family, member A1; DNA segment, Chr 18,	
S78217_s_at	protein phosphatase 1, catalytic subunit, gamma isoform	protein phosphatase type 2A activity protein phosphatase type 1 activity phosphoprotein phosphatase activity protein serine/threonine phosphatase activity, calcium-dependent protein magnesium-dependent protein
U36482_g_at	endoplasmic retuclum protein 29	
U56839_at	purinergic receptor P2Y, G-protein coupled 2	rhodopsin-like receptor activity G-protein coupled receptor activity ATP binding, purinergic nucleotide receptor activity, G-protein coupled
E01884_at	Interleukin-1 beta	cytokine activity, interleukin-1 receptor binding, growth factor activity
U75917_g_at	clathrin-associated protein 17	
X06150cds_at	glycine N-methyltransferase	catalytic activity, folic acid binding methyltransferase activity, S-adenosylmethionine-dependent methyltransferase activity, glycine N-methyltransferase activity
X06357cds_s_at	alanine-glyoxylate aminotransferase	serine-pyruvate transaminase activity alanine-glyoxylate transaminase activity
X13044_at	CD74 antigen (invariant polpypeptide of major histocompatibility class II antigen-associated)	chaperone activity
X53363cds_s_at	calreticulin	chaperone activity, calcium ion binding calcium ion storage activity
X62295cds_s_at	angiotensin II receptor, type 1 (AT1A)	rhodopsin-like receptor activity receptor activity, G-protein coupled receptor activity, angiotensin type II receptor activity

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 2 – Continued

X69834_at	serine protease inhibitor 2.4	
X81448cds_at	Similar to cytokeratin (LOC294853)	
X84047cds_at	GNAS complex locus, XLas protein	signal transducer activity,GTP binding
X86178	alpha-2-glycoprotein 1, zinc	zinc ion binding
Z50052_at	complement component 4 binding protein, pseudogene 1	
Z50144_at	aminoadipate aminotransferase	transaminase activity

A3

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4

Probe Set ID	Gene Name	Molecular Function Description
AB000491_at	proteasomal ATPase (SUG1)	
AB005547_at	aquaporin 8	transporter activity, water channel activity
AF012714_at	multiple inositol polyphosphate histidine phosphatase 1	inositol-1,3,4,5,6-pentakisphosphate 3-phosphatase activity, inositol-1,4,5,6-tetrakisphosphate 6-phosphatase activity
AF029240_at	MHC class Ib RT1.S3	
AF061971_at	palmitoyl-protein thioesterase 2	catalytic activity, palmitoyl-(protein) hydrolase activity
AF089825_at	activin beta E	hormone activity, growth factor activity
AJ000347_g_at	3(2),5-bisphosphate nucleotidase	inositol or phosphatidylinositol phosphatase activity 3'(2'),5'-bisphosphate nucleotidase activity
D00569_at	2,4-dienoyl CoA reductase 1, mitochondrial	2,4-dienoyl-CoA reductase (NADPH) activity oxidoreductase activity
D00698_s_at	insulin-like growth factor 1	hormone activity, growth factor activity
D00729_g_at	dodecenoyl-coenzyme A delta isomerase	catalytic activity, dodecenoyl-CoA delta-isomerase
D00753_at	Serine protease inhibitor	
D10874_at	ATPase, H+ transporting, lysosomal (vacuolar proton pump) 16 kDa	ATPase activity, coupled to transmembrane movement of ions
D13623_g_at	Similar to ribosome-binding protein p34 - rat (LOC287633),	
D13907_g_at	mitochondrial processing peptidase beta	metalloendopeptidase activity mitochondrial processing peptidase, hydrolase activity
D21215cds_s_	coagulation factor X	serine-type endopeptidase activity, chymotrypsin activity, trypsin activity, peptidase activity, hydrolase
D30647_at	acyl-Coenzyme A	acyl-CoA dehydrogenase activity
D30649mRN	ectonucleotide	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4 – Continued

D21215cds_s_	coagulation factor X	serine-type endopeptidase activity, chymotrypsin activity, trypsin activity, peptidase activity, hydrolase
D30647_at	acyl-Coenzyme A	acyl-CoA dehydrogenase activity
D30649mRNA_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 3	
D45254_g_at	cellular nucleic acid binding protein	single-stranded DNA binding, single-stranded RNA binding
D88250_at	complement component 1, s subcomponent	
D90211_s_at	lysosomal membrane glycoprotein 2	tRNA ligase activity, ATP binding
D90265_s_at	proteasome (prosome, macropain) subunit, alpha type 1	endopeptidase activity transporter activity peptidase activity hydrolase activity
D90404_at	cathepsin C	cysteine-type endopeptidase activity dipeptidyl-peptidase I activity, hydrolase activity
E01184cds_s_at	cytochrome P450, family 1, subfamily a, polypeptide 2	monooxygenase activity oxidoreductase activity
J00728cds_f_at	cytochrome P450, 2b19	monooxygenase activity
J04171_at	glutamate oxaloacetate transaminase 1	aspartate transaminase activity, transaminase activity
J05035_at	steroid 5 alpha-reductase 1	3-oxo-5-alpha-steroid 4-dehydrogenase activity oxidoreductase activity
J05470_at	carnitine palmitoyltransferase 2	carnitine O-palmitoyltransferase activity acyltransferase activity, transferase activity
L07736_at	carnitine palmitoyltransferase 1, liver	carnitine O-palmitoyltransferase activity acyltransferase activity, transferase activity

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4 – Continued

L13619_at	insulin induced gene 1	
M16235_at	lipase, hepatic	catalytic activity, triacylglycerol lipase activity lipid transporter activity, heparin binding, hydrolase
M22359	alpha(1)-inhibitor 3, variant I	protease inhibitor activity, inflammatory response
M32016_at	lysosomal membrane glycoprotein 2	tRNA ligase activity, ATP binding
M33648_at	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2	hydroxymethylglutaryl-CoA synthase activity transferase activity
M35270	alanine-glyoxylate aminotransferase	serine-pyruvate transaminase activity alanine-glyoxylate transaminase activity
M55015cds	nucleolin	nucleic acid binding, DNA binding, RNA binding
M64733mRNA_s_at	clusterin	response to stress, spermatogenesis
M67465_at	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase	steroid dehydrogenase activity
M74067_at	claudin 3	structural molecule activity
M81183Exon_UTR_g_at	Rat insulin-like growth factor I mRNA, 3' end of mRNA	
M81687_at	syndecan 2	cytoskeletal protein binding
M83143_g_at	sialyltransferase 1	beta-galactoside alpha-2,6-sialyltransferase activity sialyltransferase activity, transferring glycosyl groups
M86912exon_	angiotensin II receptor, type 1 (AT1A)	rhodopsin-like receptor activity, G-protein coupled receptor activity, angiotensin type II receptor activity
rc_AA799452	transaldolase 1	transaldolase activity, transferase activity
rc_AA79952	Similar to E25B protein	
rc_AA799762_g_at	Similar to RIKEN cDNA 2700038C09 (LOC296470),	
rc_AA800250	succinate dehydrogenase	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4 – Continued

rc_AA800268	Similar to Ormdl2 protein	
rc_AA800272	Similar to mitochondrial ribosomal protein L3 (L3mt)	
rc_AA800296	Similar to poly(A) polymerase	
rc_AA800318	serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor)	
rc_AA859700_g_at	Similar to Protoporphyrinogen oxidase (LOC289219), mRNA	
rc_AA891800_at	Similar to RIKEN cDNA 1110013G13 (LOC310856),	
rc_AA891914_at	aminoacylase 1	aminoacylase activity, metallopeptidase activity hydrolase activity
rc_AA891916_g_at	membrane interacting protein of RGS16	glycerophosphodiester phosphodiesterase activity
rc_AA892128_at	peroxisomal biogenesis factor 11A	
rc_AA892470_at	Similar to histone H2A.F/Z variant isoform 1; purine-rich binding element protein B (LOC314176), mRNA	
rc_AA892821_g_at	aldo-keto reductase family 7, member A2 (aflatoxin aldehyde reductase)	aldehyde reductase activity aldo-keto reductase activity
rc_AA892828_at	Similar to pyruvate dehydrogenase (lipoamide) beta (LOC289950), mRNA	
rc_AA892888_g_at	Similar to RIKEN cDNA 0610006F02 (LOC366792)	
rc_AA892916	Ab2-305 mRNA, complete cds	
rc_AA893235	Similar to G0S2-like protein	
rc_AA893353	Similar to Aspartyl	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4 – Continued

rc_AA893515_at	Similar to Ab2-292 (LOC294912), mRNA	
rc_AA893552	serine (or cysteine) proteinase)	
rc_AA893658	Similar to alcohol dehydrogenase PAN1B-like protein (LOC305150), mRNA	
rc_AA893690	Similar to neuronal protein 15.6 (LOC299310), mRNA	
rc_AA894101	Similar to N-terminal asparagine amidohydrolase	
rc_AA894130	Unknown (protein for MGC:72638)	
rc_AA894277_at	Similar to RIKEN cDNA 2010107G23 (LOC294499)	
rc_AA944007	Nucleobindin	structural constituent of bone
rc_AI103957_	CD 81 antigen	protein binding, response to wounding regulation of growth / cell proliferation
rc_AI171090_g_at	3-hydroxy-3-methylglutaryl CoA lyase	catalytic activity, hydroxymethylglutaryl-CoA lyase activity
rc_AI172476_	TGFB inducible early growth response	
rc_AI176052_at	adenylate kinase 3	adenylate kinase activity
rc_AI177161_	NF-E2-related factor 2	DNA binding
rc_AI178135_at	complement component 1, q subcomponent binding protein	
rc_AI639418_at	deiodinase, iodothyronine, type I	thyroxine 5'-deiodinase activity, selenium binding oxidoreductase activity, hydrolase activity
rc_H31813_at	Similar to cDNA sequence BC021917 (LOC361730)	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4 – Continued

S45663_g_at	glycoprotein, synaptic 2	
S61865_s_at	syndecan 1	cytoskeletal protein binding
S61960_s_at	Similar to kynurenine aminotransferase/glutamine	
S63521_i_at	heat shock 70kD protein 5	heat shock protein activity, ATP binding
S69874_s_at	fatty acid binding protein 5, epidermal	transporter activity, binding, fatty acid binding lipid binding
S74351_s_at	dual specificity phosphatase 1	MAP kinase phosphatase activity
U05675_at	fibrinogen, B beta polypeptide	blood coagulation, wound healing
U08976_at	enoyl coenzyme A hydratase 1, peroxisomal	catalytic activity, isomerase activity
U28504_g_at	solute carrier family 17 (sodium phosphate), member 1	transporter activity, sodium:phosphate symporter activity, phosphate transporter activity, symporter
U61729_at	proline rich 2	
U66322_at	leukotriene B4 12-hydroxydehydrogenase	response to toxin
U89280_at	hydroxysteroid (17-beta) dehydrogenase 9	
X04070_at	gap junction membrane channel protein beta 1	gap-junction forming channel activity connexon channel activity
X05684_at	pyruvate kinase, liver and RBC	magnesium ion binding, pyruvate kinase activity kinase activity, transferase activity
X06107_r_at	insulin-like growth factor 1	hormone activity, growth factor activity
X51615_g_at	Hypothetical LOC290275	
X54096_at	lecithin cholesterol acyltransferase	catalytic activity, phosphatidylcholine-sterol O-acyltransferase activity, transferase activity
X62660mRNA_g_at	Similar to GLUTATHIONE S-TRANSFERASE 8 (GST 8-8) (CHAIN 8) (GST CLASS-	

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 4 – Continued

X79807_at	coagulation factor X	serine-type endopeptidase activity, chymotrypsin activity, trypsin activity, peptidase activity, hydrolase
X86561cds#2	fibrinogen, alpha polypeptide	blood coagulation tissue regeneration
X91234_at	hydroxysteroid (17-beta) dehydrogenase 2	estradiol 17-beta-dehydrogenase activity oxidoreductase activity
X97443_s_at	integral membrane protein Tmp21-I (p23)	protein carrier activity, protein transporter activity
Z50144_g_at	aminoadipate aminotransferase	transaminase activity

A4

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 5

Probe Set ID	Gene Name	Molecular Function Description
AF023087	early growth response 1	nucleic acid binding, DNA binding transcription factor activity, transcriptional
rc_AA818593	phosphatidate phosphohydrolase type 2a	phosphatidate phosphatase activity
rc_AA892498_at	Similar to tetraspanin TM4-A homolog (LOC300733), mRNA	
rc_AI012030_at	matrix Gla protein	calcium ion binding
rc_AI232379_at	platelet derived growth factor receptor, alpha polypeptide	protein kinase activity, protein serine/threonine kinase activity, protein-tyrosine kinase activity transmembrane receptor protein tyrosine kinase
rc_AI639058_s_at	Similar to Nedd4 WW binding protein 4 (LOC311676), mRNA	
U90261UTR#1_	SH3-domain kinase binding	protein binding, regulation of apoptosis
V01227_s_at	tubulin, alpha 1	structural molecule activity, GTP binding protein domain specific binding, protein heterodimerization activity, microtubule-based
X06916_at	S100 calcium-binding protein A4	calcium ion binding

A5

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 6

Probe Set ID	Gene Name	Molecular Function Description
D12769_at	basic transcription element binding protein 1	
D32249_s_at	praja 2, RING-H2 motif containing	neurogenesis
D89655_at	scavenger receptor class B, member	
J03190_g_at	aminolevulinic acid synthase 1	5-aminolevulinate synthase activity, acyltransferase transaminase activity, heme biosynthesis
L00191cds#1_s_at	fibronectin 1	heparin binding, oxidoreductase activity mercury ion binding, acute-phase response cell adhesion, metabolism, wound healing
L09647_at	forkhead box A2	DNA binding, transcription factor activity
L28135_at	solute carrier family 2 (facilitated glucose transporter), member 2	transporter activity, sugar porter activity glucose transporter activity, carrier activity
L36532_s_at	complement receptor related protein	
M77479_at	solute carrier family 10 (sodium/bile acid cotransporter family), member 1	carrier activity, bile acid:sodium symporter activity symporter activity
rc_AA799729_g	phosphodiesterase 4B	cAMP-specific phosphodiesterase activity
rc_AA799803_at	Similar to complement component 1, r subcomponent (LOC312705)	
rc_AA859869_s_at	proteasome (prosome, macropain) 26S subunit, non-ATPase, 1	
rc_AA859954_a	vacuole Membrane Protein 1	
rc_AA874784_s	lipase A, lysosomal acid	catalytic activity, sterol esterase activity, hydrolase activity
rc_AA875097_a	fibrinogen, alpha polypeptide	blood coagulation, tissue regeneration
rc_AA891194_s	Arg/Abl-interacting protein ArgBP2	intracellular signaling cascade
rc_AA891591_a	programmed cell death 8	DNA fragmentation during apoptosis

List of Gene Ontology Annotated Transcripts from 2-week SOM Cluster 6 - Continued

rc_AA892545_a	Similar to ITM (LOC309131)	
rc_AA892547_at	Similar to RIKEN cDNA 2010311D03 (LOC362134), mRNA	
rc_AA893239_a	2-hydroxyphytanoyl-CoA lyase	
rc_AA957510_s_at	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	magnesium ion binding, calcium-transporting ATPase activity, calcium ion binding ATP binding, ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism, hydrolase activity
S81478_s_at	dual specificity phosphatase 1	MAP kinase phosphatase activity
U02553cds_s_at	dual specificity phosphatase 1	MAP kinase phosphatase activity
U06230_s_at	protein S (alpha)	
U18762_at	retinol dehydrogenase type III	retinol dehydrogenase activity
U19614_g_at	lamina-associated polypeptide 1C	lamin binding
U32314_g_at	Pyruvate carboxylase	catalytic activity, pyruvate carboxylase activity ATP binding, biotin binding ligase activity, manganese ion binding
X55298_at	ribophorin II	dolichyl-diphosphooligosaccharide-protein glycotransferase activity, transferase activity
X78997_at	cadherin 17	calcium ion binding, cell adhesion
X95096_at	Macrophage stimulating 1 (hepatocyte growth factor-like)	thrombin activity, serine-type endopeptidase activity, chymotrypsin activity, trypsin activity calcium ion binding, peptidase activity, hydrolase
Y14933mRNA_s_at	one cut domain, family member 1	DNA binding transcription factor activity, glucose metabolism, regulation of transcription

B1

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 1

Probe Set ID	Gene Name	Molecular Function Description
AA799276_at	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	magnesium ion binding, calcium-transporting ATPase activity, calcium ion binding
AF030091UTR	cyclin L1	
AF037072_at	carbonic anhydrase 3	carbonate dehydratase activity, zinc ion binding lyase activity
AF063102_at	latrophilin 2	G-protein coupled receptor activity
AF077000_at	protein tyrosine phosphatase, non-receptor type 23	
AF079864_at	olfactory receptor gene Olr59	olfactory receptor activity
AF089825_at	activin beta E	hormone activity, growth factor activity
D00698_s_at	insulin-like growth factor 1	hormone activity growth factor activity
D28557_s_at	cold shock domain protein A	DNA binding
J00738_s_at	alpha-2u globulin PGCL4	
J03865mRNA_f_a		Terpenoid biosynthesis
L38482_at	Similar to cell division cycle 34; ubiquitin-conjugating enzyme	
rc_AA800029_at	Similar to highwire; PAM; rpm	
rc_AA875097_at	fibrinogen, alpha polypeptide	Blood clotting cascade
rc_AI008815_s_at	Similar to Cytochrome c, somatic (LOC365738), mRNA	Apoptosis, Electron-transport chain
X13058_at	tumor suppressor protein p53	Apoptosis, DNA binding, transcription factor
rc_AI071866_s_at	Nclone10 mRNA	
rc_AI236721_r_at		Cell-Cycle
L14680_at	B-cell lymphoma 2	Protein binding, cell survival
U33500_g_at	retinol dehydrogenase type II	

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 1 - Continued

U37138_at	steroid sulfatase	steryl-sulfatase activity, sulfuric ester hydrolase
X06107_r_at	insulin-like growth factor 1	hormone activity, growth factor activity
X13119cds_s_at		Terpenoid Biosynthesis
X14552_at	alpha-2u globulin PGCL4	
X53588_at	glucokinase	glucokinase activity, hexokinase activity, ATP
X84210complete	nuclear factor I/A	DNA binding transcription factor activity
X86561cds#2_at	fibrinogen, alpha polypeptide	
Y14933mRNA	one cut domain, family	DNA binding, transcription factor activity

B2

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 2

Probe Set ID	Gene Name	Molecular Function Description
AA801441_at	platelet-activating factor acetylhydrolase, isoform Ib	1-alkyl-2-acetylglycerophosphocholine esterase activity, hydrolase activity
AB009890_at	trophoblast specific protein	
AF034582_g_at	SEC31-like 1 (S. cerevisiae)	
D12769_at	basic transcription element binding protein 1	
J05035_g_at	steroid 5 alpha-reductase 1	3-oxo-5-alpha-steroid 4-dehydrogenase activity oxidoreductase activity
L23413_at	solute carrier family 26 (sulfate transporter),member 1	sulfate porter activity, sulfate transporter activity, antiporter activity
M27440_at	apolipoprotein B	
M58758_g_at	ATPase, H+ transporting	hydrogen ion transporter activity
rc_AA817854_s_a	ceruloplasmin	Ferroxidase activity, copper ion binding
rc_AI105348_i_at	cofilin 1	actin binding
rc_AI171355_s_at	cytochrome b, mitochondrial	electron transporter activity
rc_AI171630_s_at	mitogen activated protein kinase 14	protein serine/threonine kinase activity protein kinase CK2 activity, MAP Kinase activity cAMP-dependent protein kinase activity
rc_AI180108_at	Similar to Wbscr1	
U58858_at	junction plakoglobin	
U66723_s_at	solute carrier family 28	nucleoside:sodium symporter activity
X52140_at	integrin alpha 1	magnesium ion binding receptor activity
X55660_at	proprotein convertase subtilisin/kexin type3	

B3

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 3

Probe Set ID	Gene Name	Molecular Function Description
AB017544	peroxisomal biogenesis factor 14	peroxisome organization and biogenesis
D13417_g_at	hairy and enhancer of split 1	DNA binding, transcriptional repressor activity
D86745cds_s_at	nuclear receptor subfamily 0, group B, member 2	ligand-dependent nuclear receptor activity
D89375_s_at	dopa/tyrosine sulfotransferase	
J03190_at	aminolevulinic acid synthase 1	5-aminolevulinate synthase activity, acyltransferase activity, transaminase activity, transferase activity
M93257_s_at	catechol-O-methyltransferase	magnesium ion binding, methyltransferase activity O-methyltransferase activity, S-adenosylmethionine-dependent methyltransferase
M96548_at	zinc finger protein 354A	nucleic acid binding, DNA binding, transcription factor activity
rc_AA800005_a	CD151 antigen	
rc_AA892775	lysozyme	
rc_AA893239_a	2-hydroxyphytanoyl-CoA lyase	
rc_AA893384_	Similar to interferon regulatory factor 3 (LOC292892), mRNA	
rc_AA946040_at	Similar to Uroplakin Ia (UPIa) (UPKa) (LOC365227), mRNA	
rc_AI011998_at	DnaJ (Hsp40) homolog, subfamily B, member 9	chaperone activity
rc_AI171085_at	ribosomal protein L39	RNA binding, structural constituent of ribosome
rc_AI175935_at	myosin IE	motor activity, ATP binding
rc_AI180442_at	farensyl diphosphate synthase	Dimethylallyltranstransferase, farnesyltranstransferase activity geranyltranstransferase activity
rc_AI230260_s_	casein kinase II beta subunit	protein serine/threonine kinase activity

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 3 - Continued

rc_AI235747_at	Glutathione-S-transferase, a2	glutathione transferase activity, transferase activity
rc_AI639418_at	deiodinase, iodothyronine, typeI	thyroxine 5'-deiodinase activity, selenium binding oxidoreductase activity, hydrolase activity
X15580	6-phosphofructo-2-kinase/fructose-2,6 biphosphatase 1	catalytic activity, 6-phosphofructo-2-kinase activity fructose-2,6-bisphosphate 2-phosphatase activity ATP binding, kinase activity, transferase activity hydrolase activity
X55660_g_at	proprotein convertase subtilisin/kexin type3	
X60769mRNA_	CCAAT/enhancer binding protein (C/EBP), beta	DNA binding, protein binding transcriptional activator activity

B4

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 4

Probe Set ID	Gene Name	Molecular Function Description
J02612mRNA	UDP glycosyltransferase 1 family polypeptide, A1-A8	glucuronosyltransferase activity transferring glycosyl groups, hexosyl
L21711_s_at	lectin, galactose binding	Sugar binding
M25804_g_at	nuclear receptor subfamily 1, group D, member 1	DNA binding, transcription factor activity steroid hormone receptor activity receptor
M27467_at	cytochrome oxidase subunit VIc	
M33025_s_at	parathymosin	enzyme inhibitor activity, zinc ion binding
X05684_at	pyruvate kinase, liver and RBC	magnesium ion binding, pyruvate kinase

B5

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 5

Probe Set ID	Gene Name	Molecular Function Description
AF014503_at	nuclear protein 1	transcription factor activity
AF045464_s_at	aldo-keto reductase family 7, member A3 (aflatoxin aldehyde reductase)	aldehyde reductase activity, aflatoxin metabolism
AJ011656cds_s_at	claudin 3	structural molecule activity
D28339_s_at	3-hydroxyanthranilate 3,4-dioxygenase	3-hydroxyanthranilate 3,4-dioxygenase activity oxidoreductase activity
D83796_s_at	UDP glycosyltransferase 1 family polypeptide, A1, A2, A4, A6, A7, A8	glucuronosyltransferase activity transferring glycosyl group, transferring hexosyl groups, bilirubin conjugation metabolism
E01184cds_s_at	cytochrome P450, family 1, subfamily a, polypeptide 2	monooxygenase activity, oxidoreductase activity
E01524cds_s_at	P450 (cytochrome), oxidoreductase	NADPH-hemoprotein reductase activity electron transporter activity, oxidoreductase
E12625cds_at	sterol-C4-methyl oxidase	catalytic activity, oxidoreductase activity
J03179_g_at	D site albumin promoter	DNA binding, transcriptional activator
J05210_g_at	ATP citrate lyase	magnesium ion binding, catalytic activity ATP citrate synthase activity, lyase activity ligase activity
K01934mRNA	thyroid hormone responsive	
K02815_s_at	RT1 class II, locus Ba	
L00320cds_f	cytochrome P450, 2b19	monooxygenase activity
L38615_g_at	glutathione synthetase	glutathione synthase activity, ATP binding ligase activity
M10068mRNA	P450 (cytochrome) oxidoreductase	NADPH-hemoprotein reductase activity electron transporter activity oxidoreductase activity

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 5 - Continued

M21476_s_at	prolyl 4-hydroxylase, beta polypeptide	protein disulfide isomerase activity procollagen-proline 4-dioxygenase activity electron transporter activity
M26127_s_at	cytochrome P450, family 1, subfamily a, polypeptide 2	monooxygenase activity oxidoreductase activity
M74067_at	claudin 3	structural molecule activity
M89945mRNA_g_at	farensyl diphosphate synthase	dimethylallyltranstransferase activity farnesyltranstransferase activity geranyltranstransferase activity transferase activity
rc_AA799778_at	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit b, isoform 1	hydrogen ion transporter activity
rc_AA852004	glutamine synthetase 1	glutamate-ammonia ligase activity
rc_AA892828	Similar to pyruvate dehydrogenase (lipoamide) beta (LOC289950), mRNA	
rc_AA945050	rat senescence marker protein 2A gene, exons 1 and 2	
rc_AI008131_s_at	S-adenosylmethionine decarboxylase 1	adenosylmethionine decarboxylase activity lyase activity, carboxy-lyase activity
rc_AI030175_s_at	sorbitol dehydrogenase	L-iditol 2-dehydrogenase activity alcohol dehydrogenase activity, zinc-dependent, zinc ion binding oxidoreductase activity
rc_AI137856_s_at	P450 (cytochrome) oxidoreductase	NADPH-hemoprotein reductase activity electron transporter activity, oxidoreductase
rc_AI169695_f_at	MRNA for hydroxysteroid sulfotransferase subunit,	
rc_AI171734_s_	fumarate hydratase 1	catalytic activity, fumarate hydratase lyase activity

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 5 - Continued

rc_AI175764_s_	stearoyl-Coenzyme A desaturase 1	
rc_AI176052_at	adenylate kinase 3	adenylate kinase activity
rc_AI230712_at	Subtilisin - like endoprotease	serine-type endopeptidase activity subtilase activity, transmembrane receptor protein tyrosine kinase activity, electron transporter activity, ATP binding peptidase activity, serine-type peptidase activity hydrolase activity
rc_AI232691_at	lectin, galactose binding, soluble 8	sugar binding
rc_AI235585_s_ at	cathepsin D	aspartic-type endopeptidase activity cathepsin D activity, pepsin A activity hydrolase activity
rc_H31813_at	Similar to cDNA sequence BC021917 (LOC361730),	
S56937_s_at	UDP glycosyltransferase 1 family polypeptide A1, A2, A4, A6, A7, A8	glucuronosyltransferase activity, transferring glycosyl groups, transferring hexosyl groups
U10357_at	pyruvate dehydrogenase kinase, isoenzyme 2	[pyruvate dehydrogenase (lipoamide)] kinase activity, ATP binding kinase activity transferase activity
U17971_at	protein tyrosine phosphatase 2E	protein tyrosine phosphatase activity
U67915_at	mast cell protease 1	serine-type endopeptidase activity chymase activity
X13044_at	CD74 antigen (invariant polpypeptide of major histocompatibility class II antigen-associated)	chaperone activity
X59859_i_at	decorin	collagen binding, protein N-terminus binding
X70871_at	cyclin G1	cyclin-dependent protein kinase regulator
X72792cds_s_at	alcohol dehydrogenase 1	alcohol dehydrogenase activity, zinc ion binding, oxidoreductase activity
X91234_at	hydroxysteroid (17-beta) dehydrogenase 2	estradiol 17-beta-dehydrogenase activity oxidoreductase activity
Z15123exon#5_ s_at	S-adenosylmethionine decarboxylase 1	adenosylmethionine decarboxylase activity lyase activity, carboxy-lyase activity

B6

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 6

Probe Set ID	Gene Name	Molecular Function Description
AB002558_at	glycerol-3-phosphate dehydrogenase 1 (soluble)	glycerol-3-phosphate dehydrogenase (NAD+) activity
AF001898_at	aldehyde dehydrogenase family 1, member A1	aldehyde dehydrogenase, oxidoreductase activity
AF003835_at	isopentenyl-diphosphate delta isomerase	magnesium ion binding, isopentenyl-diphosphate delta-isomerase activity
AF036761_g_at	stearoyl-Coenzyme A desaturase 1, stearoyl-Coenzyme A desaturase 2	
AF061971_at	palmitoyl-protein thioesterase 2	catalytic activity, palmitoyl-(protein) hydrolase activity, hydrolase activity
AF062389_at	kidney-specific protein (KS)	
AF072411_at	cd36 antigen	receptor activity, fatty acid binding
D10655_at	dihydrolipoamide acetyltransferase	nucleotide binding, dihydrolipoyllysine-residue acetyltransferase activity, protein binding dihydrolipoamide S-acyltransferase activity
D37920_at	squalene epoxidase	monooxygenase activity, oxidoreductase activity
E00717UTR#1_s_at	cytochrome P450, family 1, subfamily a, polypeptide 1	monooxygenase activity, oxidoreductase activity
E00778cds_s_at	cytochrome P450, family 1, subfamily a, polypeptide 1	monooxygenase activity, oxidoreductase activity
E03358cds_at	proteasome (prosome, macropain) subunit, alpha type 2	endopeptidase activity, peptidase activity hydrolase activity
J02679_s_at	NAD(P)H dehydrogenase, quinone 1	NAD(P)H dehydrogenase (quinone) activity oxidoreductase activity
J05460_s_at	cytochrome P450, family 7, subfamily a, polypeptide 1	monooxygenase activity, cholesterol 7-alpha-monooxygenase activity, oxidoreductase activity
K03241cds_s_at	cytochrome P450, family 1, subfamily a, polypeptide 2	monooxygenase activity, oxidoreductase activity

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 6 - Continued

L12380_at	ADP-ribosylation factor 1	GTPase activity, receptor signaling protein, GTP binding, protein transporter activity
L13619_at	insulin induced gene 1	
L22294_at	pyruvate dehydrogenase kinase 1	[pyruvate dehydrogenase (lipoamide)] kinase ATP binding, kinase activity, transferase activity
M26594_at	malic enzyme 1	malic enzyme activity, malate dehydrogenase (decarboxylating) activity, malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) activity oxidoreductase activity
M33329_f_at	rat senescence marker protein 2A gene, exons 1 and 2	
M36151cds_s_at	RT1 class II, locus Bb	
M37584_at	CDNA clone MGC:72814 IMAGE:6922321, complete cds	
M63983_s_at	hypoxanthine guanine phosphoribosyl transferase	magnesium ion binding hypoxanthine phosphoribosyltransferase activity transferase activity, transferring glycosyl groups
rc_AA799279	Similar to mitochondrial carrier homolog 2	
rc_AA799326	cd36 antigen	receptor activity, fatty acid binding
rc_AA799442_at	Hypothetical LOC293114 (LOC293114), mRNA	
rc_AA799442_g_at	Hypothetical LOC293114 (LOC293114), mRNA	
rc_AA800053_at	Similar to androgen receptor N-terminal-interacting protein (LOC289508), mRNA	
rc_AA800243_at	Similar to cell death activator CIDE-A (LOC291541), mRNA	
rc_AA801174_at	Rat ig delta heavy chain	

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 6 - Continued

rc_AA859980	Similar to acetyl CoA transferase-like (LOC308100), mRNA	
rc_AA875126_g_at	Similar to Myosin Id (Myosin heavy chain myr 4) (LOC289785), mRNA	
rc_AA891226	proteasome (prosome, macropain) subunit, beta type 5	endopeptidase activity, peptidase activity hydrolase activity
rc_AA891950_at	Similar to RIKEN cDNA 1810021J13 (LOC300516), mRNA	
rc_AA892561_at	Similar to RIKEN cDNA 4931406C07 (LOC363016), mRNA	
rc_AA892801_at	eukaryotic translation elongation factor 2	translation elongation factor activity, ATP binding, GTP binding
rc_AA892835_at	Similar to RIKEN cDNA 5730434I03 gene	
rc_AA892842_at	Similar to capping protein alpha 2 subunit (LOC368086),	
rc_AA892950_at	Similar to mitochondrial DNA polymerase accessory subunit	
rc_AA893173_at	Similar to vacuolar protein sorting 29 isoform 2	
rc_AA925752_at	cd36 antigen	receptor activity, fatty acid binding
rc_AA944856_at	RAP1B, member of RAS oncogene family	GTPase activity, GTP binding
rc_AA945082_at	glutathione-S-transferase a2	glutathione transferase activity
rc_AA946368_at	cd36 antigen	receptor activity, fatty acid binding
rc_AA996484_g_	aldehyde dehydrogenase	aldehyde dehydrogenase activity, oxidoreductase

List of Gene Ontology Annotated Transcripts from 12-week SOM Cluster 6 - Continued

rc_AA997614_s_at	cytochrome P450, subfamily 51	trypsin activity, monooxygenase activity methyltransferase activity, sterol 14-demethylase activity, oxidoreductase activity
rc_AI171506_g_at	malic enzyme 1	malic enzyme activity, malate dehydrogenase (decarboxylating) activity, malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) activity oxidoreductase activity
rc_AI175486_at	ribosomal protein S7	RNA binding, structural constituent of ribosome
S35751_f_at	3-alpha-hydroxysteroid dehydrogenase	
S69874_s_at	fatty acid binding protein 5, epidermal	transporter activity binding, fatty acid binding lipid binding
S72506_s_at	glutathione-S-transferase, alpha type2	glutathione transferase activity transferase activity
S82820mRNA_s_	glutathione-S-transferase, alpha type2	glutathione transferase activity transferase activity
U23056_at	CEA-related cell adhesion molecule 10	
U64705cds_i_at	Similar to translation initiation factor eIF-4A II -	
X12535cds_at	RAB7, member RAS oncogene family	DNA binding, GTPase activity, ATP binding GTP binding, protein transporter activity
X53054_g_at	RT1 class II, locus Bb	
X61654_at	CD63 antigen	
X63410cds_f_at	rat senescence marker protein 2A gene, exons 1 and 2	
X68782cds_at	Rat ig delta heavy chain constant region and 3' ut, mrna	
X77235_at	ADP-ribosylation-like 4	GTP binding
X92097_at	coated vesicle membrane	protein carrier activity, protein transporter

C

Gene Ontology Molecular Function Classification of Transcripts.

		2-weeks			12-weeks	
Rat Transcripts	**Genbank**	**Dose (mg/g)**			**Dose (mg/g)**	
Gene Family	**Accession #**	**0.01**	**0.1**	**0.01**	**0.1**	**1**
Cytochrome P450 Enzymes						
CYPC17 Hydroxylase	M21208					
CYPXIB2	D11354					
CYP2D18	U48220					
cYP11B2 gene(aldosterone synthase)	D14097					
CYP1A1	E01184					
Cytochrome P450 15-beta	J03786					
CYP1A2	E00778					
CYP450 1B	rc_AI176856					
CYP450	E00717					
Transcription Regulators						
Mitogen-Activated Protein Kinase Kinase Kinase	M94454					
Protein Phosphatase 3, Catalytic subunit, beta isoform	M31809					
Protein Phosphatase 3, (regulating subunit B, alpha	D10393					
Casein Kinase1, alpha 1	rc_AI235291					
Casein kinase II, alpha 1 polypeptide	rc_AA875506					
Tsx	X99797					
Protein Kinase C	D85435					
Protein Tyrosine Phosphatase	U69673					
Heat shock protein	S75280					
Tyrosine Kinase receptor	L26525					
ras-related protein p23	X12535					
Jun-D	D26307					
bcl-2	L14680					
c-myc	Y00396					
Aldehyde dehydrogenase	AF001898					
Interferon Gamma Receptor	U68272					
NF1-like DNA binding protein	AB012230					
Phosphoryl kinase	M98826					
Prostatic Acid Phosphatase	M32397					
HNF-3/Forkhead homolog	L13201					
Stat 5	U24175					
Enigma homolog	U48247					
Cyt. B5	rc_AA817685					
PACAP	D14909					
G-Protein Coupled Receptor 6	AF064706					
Bradykinin receptor B1	AJ132230					
Protein Kinase , AMP-activated, alpha 1 catalytic	U40819					
myc box dependent interacting protein	rc_AI102031					

Gene Ontology Molecular Function Classification of Transcripts - Continued

Peptidyglycine alpha amidating monoxygenase	M82845						
Myogenic Regulatory fact. (MYOD)	M84176						
LIM-motif containing protein kinase 1	D31873						
Meprin 1 alpha	S43408						
Janus Kinase 2 (Protein Tyrosine kinase)	U13396						
Protein Tyrosine Phosphatase, non-receptor type 1	M33962						
protein Tyrosine Phosphatase, receptor type 2	U09357						
Transcription elongation factor A2	D12927						
BHF-1	D82074						
Olf-1	L24051						
T16 mRNA	D89730						
pericentriolar material	U95920						
Phosphoribosylpurophosphate synthetase II	X16555						
Cox-2	S67722						
High sulfur Protein B2E	AB003753						
MIPP65	AB000098						
Ras p21 Small GTP binding	M75153						
APOB	M21842						
Enzyme Regulators							
Tissue inhibitor of metalloproteinase I	rc_AI169327						
Biosynthesis & Metabolism							
L-3-Hydroxyacyl-CoA dehydrogenase, short chain	rc_AA891362						
HMG-COA reductase	X55286						
Phosphofructokinase	J04197						
Phospholipase A2	U03763utr						
Pyruvate Dehydrogenase phosphatase	AF062740						
Acetyl CoA Carboxylase	J03808						
Glycerol-3-phosphate dehydrogenase 2	U83880						
Glycerol-3-phosphate dehydrate dehydrogenase	AB002558						
phospholipase C beta	L15556						
3-Phosphoglycerate Dehydrogenase	X97772						
Mast Cell Protease 9	U72143						
Prostaglandin 12 Synthase	U53855						
Glutaminase	rc_AI176504						
Pancreatic Lipase	D88534						
Glutamate receptor-B	M36419						
Stearyl CoA desaturase	rc_AA875269						
Stearyl CoA	AF036761						
Glycogen synthase	J05446						
Acyl-CoA dehydrogenase	U64451						
Cyt. Oxidase V111	U40836						
Anti-Oxidants							
Glutathione-S-Transferase alpha type 2	rc_AA945082						
gluthatione-S-Transferase yc2	S82820						

Gene Ontology Molecular Function Classification of Transcripts - Continued

Glutathione peroxidase	D00680						
CU-Zn SOD	M21060						
NADPH Cytochrome P450 Reductase	E01524						
NADH Cytochrome b5 reductase	J03867						
NADPH Quinone Reductase	M58495						
Immune System							
Interleukin -6B	M26744						
IL-1 beta	E01884						
Fos-like antigen	rc_AI230842						
Interleukin -1 alpha	D00403						
IgM kappa chain	S81289						
T-cell receptor alpha chain	L37966						
T-cell receptor beta chain							
Vascular Cell Adhesion	M84488						
Platelet Derived Growth factor receptor alpha	Z14118						
Prostaglandin F2 receptor negative regulator	rc_AI145502						
MHC class II	X14254						
Cell Growth & Maintenace							
Fibroblast Growth Factor 16	AB002561						
3-Hydroxy-3-methylglutaryl Coenzyme A-reductase	X55286						
Transforming growth factor	L26110						
Hepatocyte Growth factor	D90102						
X-ray repair cross complementation	rc_AA893188						
Insulin like growth factor	M81183						
A Disintegirn and Metalloproteinase Domain 17 (Adam)	AJ012603						
Basic fibroblast growth factor	X07285						
Structural							
Crystallin, gamma D	X57169						
Fast Myosin Alkali Light Chain	L00088						
Actin, beta	rc_AA874855						
H36-alpha 7 integrin	X65036						
Angiotensin/Converting Enzyme	U03734						
Proteolipid Protein	M25888						
Unconventional Myosin from rat3	X7769						
beta tubulin T beta	X03369						
Oncogenes							
Tumor protein P53	X13058						
Rat heart derived c-ros 1 protooncogene	M35104						
Ost oncogene	Z35654						
Fyn-Protooncogene	rc_AI009191						
Electron Transport & Conjugation							
UDP-glucuronsyltransferase	D38061						
Sulfotransferase family, cytosolic 2A, DHEA	J02643						

Gene Ontology Molecular Function Classification of Transcripts - Continued

Arylamine N-acetyl transferase	U01348						
Peroxisomal							
Peroxisome forming factor	E03344						
Binding							
DNA Binding Protein	12752						
Oncomodulin	J02705						
CCAAT/enhancer binding protein	X12752						
Killer cell lectin-like receptor subfamily C	AF021349						
Guanylate binding protein 2, Interferon Inducible	M80367						
Calnexin	rc_AI235707						
Ubiquitin-Conjugating Enzyme E2D2	U56407						
Isopentenyl Diphosphate delta isomearse	rc_AI176778						
Forkhead box 01	M87634						
Synaptophysin	X06655						
Myristoylated alanine rich protein kinase c	rc_AA859896						
Oxidized Low Density Lipoprotein Receptor	AB005900						
Hypocretin Receptor	AF041244						
Opiod Receptor	D16349						
Neurogenic Differentiation	D82868						
D site albumin promoter binding protein	J03179						
Dynamin 2	L25605						
Retinol binding protein	M13947						
Neuronal d4 domain family member	X66022						
Gamma Aminobutyric Acid A receptor, rho 1	X95579						
A disintegirn and Metalloproteinase Domain 3 (Adam3)	Y07903						
Calcium Independent alpha-latrotoxin receptor homolog	AF063103						
Calpain 3	AF061726						
Sialytransferase 8 (alpha-2,8-sialytransferase)B	L13445						
Metallothionein 3	rc_AA924772						
Nuclear Receptor Subfamily 2, group F, Member 2	rc_AI012183						
Nuclear Receptor Subfamily 4, group A , Member 3	rc_AI176710						
Nuclear Receptor binding factor 1	AB015724						
Carbonic anhydrase 4	S68245						
Carbonic anhydrase 3	AF037072						
Sry-related HMG box prot. Sox II	AJ004858						
DNA Binding (N5)mRNA	L31882						
Zinc Finger Protein 35A	M96548						
Neural Zinc Finger Factor 3	u67080						
liver mrNA	U96490						
Bioenergetics & Transport							
Calcium Channel voltage dependent alpha 1A subunit	AF051526						
GABA receptor rho-3 subunit	D50671						

Gene Ontology Molecular Function Classification of Transcripts - Continued

Solute carrier family 21, member 7	AF041105						
Sodium Channel, voltage gated type 2, alpha	M22254						
Solute Carrier family 1	X63744						
Solute carrier family 8, member 1	X68191						
Potassium inwardly rectifying channel	x83585						
Glycine Receptor,alpha 1 subunit	D00833						
Vacuolar Protein Sorting homolog (r-vps33b)	rc_AI072341						
Probasin	M27156						
Potassium voltage-gated channel	J04731						
ATP-Regulated Potassium channel	AF081366						
Cell Signalling/Communication							
3-phosphoinositide dependent protein kinase -1	Y15748						
Thymopoietin (lamina associated polypeptide)	rc_AA818982						
Cyclic AMP regulatory phosphoprotein	S65091						
Drosophila discs-large tumor suppressor homologue	U14950						
Secreted Frizzled Related Protein 4	AF012891						
Prostaglandin E receptor 3	D16443						
Synapsin 2	rc_AI145494						
Ephrin A1	D38056						
Muscarinic Receptor	J03025						
Endothelial Differentiation Sphingolipid G-protein Rece	U10303						
Ca/Calmodulin dependent protein kinase 2 delta subun	rc_AA925495						
Chorionic Somatommatropin hormone 1 variant	rc_AI179075						
Histone 2B	rc_AA964055						
High Mobility group box 2	rc_AA996401						
Myelin Oligodendrocyte Glycoprotein	M99485						
CD5 antigen	D10728						
Prostaglandin F2 receptor negative regulator	D28581						
5-Hydroxytryptamine (Serotonin) receptor alpha	L22558						
Toll-like Receptor 4	rc_AI030997						
Adenylate Cyclase Activating Polypeptide	rc_AI228407						
CD36 antigen	rc_AA946368						
Fibroblast growth factor	x56551						
Inositol Triosphosphate	M64698						
Phosphatidy inositol-4-kinase	D83538						
5-Hydroxytryptamine(serotonin) receptor 7	L22568						
Adenosine A1 receptor	M64299						
Glial cell line derived neutrophic factor	U97143						
Transmembrane protein tMDC 1	Y07903						
calcium/calmodulin protein Kinase	L13407						
CAMP Kinase	M12492						
Serotonin transporter	X63253						
Hormone Related							
Thyroid Hormone Receptor, beta	J03819						
Corticotropin Releasing Hormone	M64987						
Luteinizing Hormone subunit beta	U25802						
Thyroxine binding globulin	M63991						
Androgen Repressible Liver prot	J02643						
Parathyroid Hormone Receptors	M77184						
Prolactin-like protein	rc_AI013523						
Estrogen receptor	S52128						

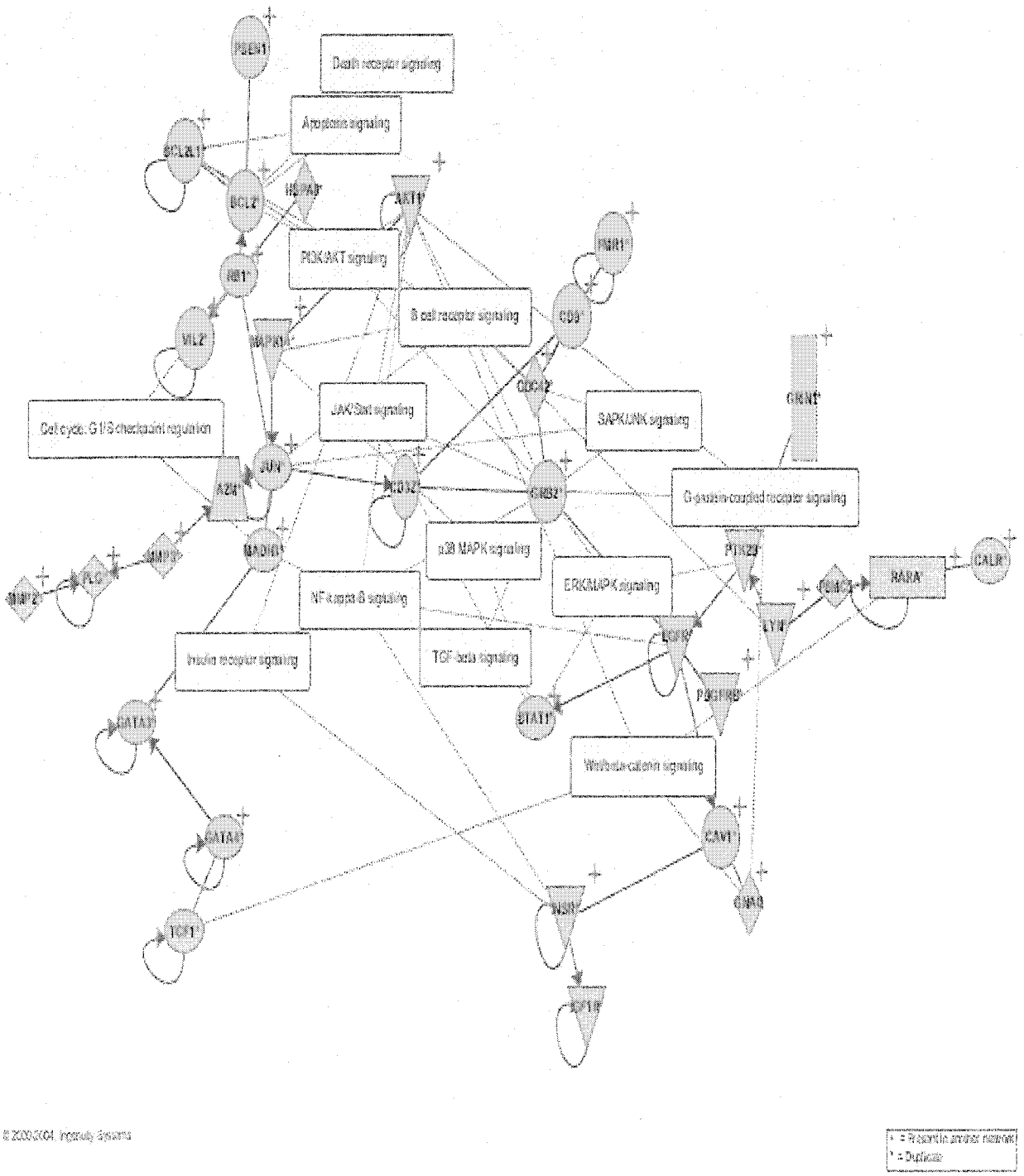
Gene Ontology Molecular Function Classification of Transcripts - Continued

Other							
Basic helix-loop helix domain	AF009329						
Defensin	U50353						
Sperm Adhesion Molecule	X89999						
Tuberous Sclerosis	U24150						
Amelogenin	U01245						
Tonin	M11564						
RVLG(vasa-like gene protein)	S75275						

Transcripts that increase or decrease are indicated by light and dark gray, respectively.

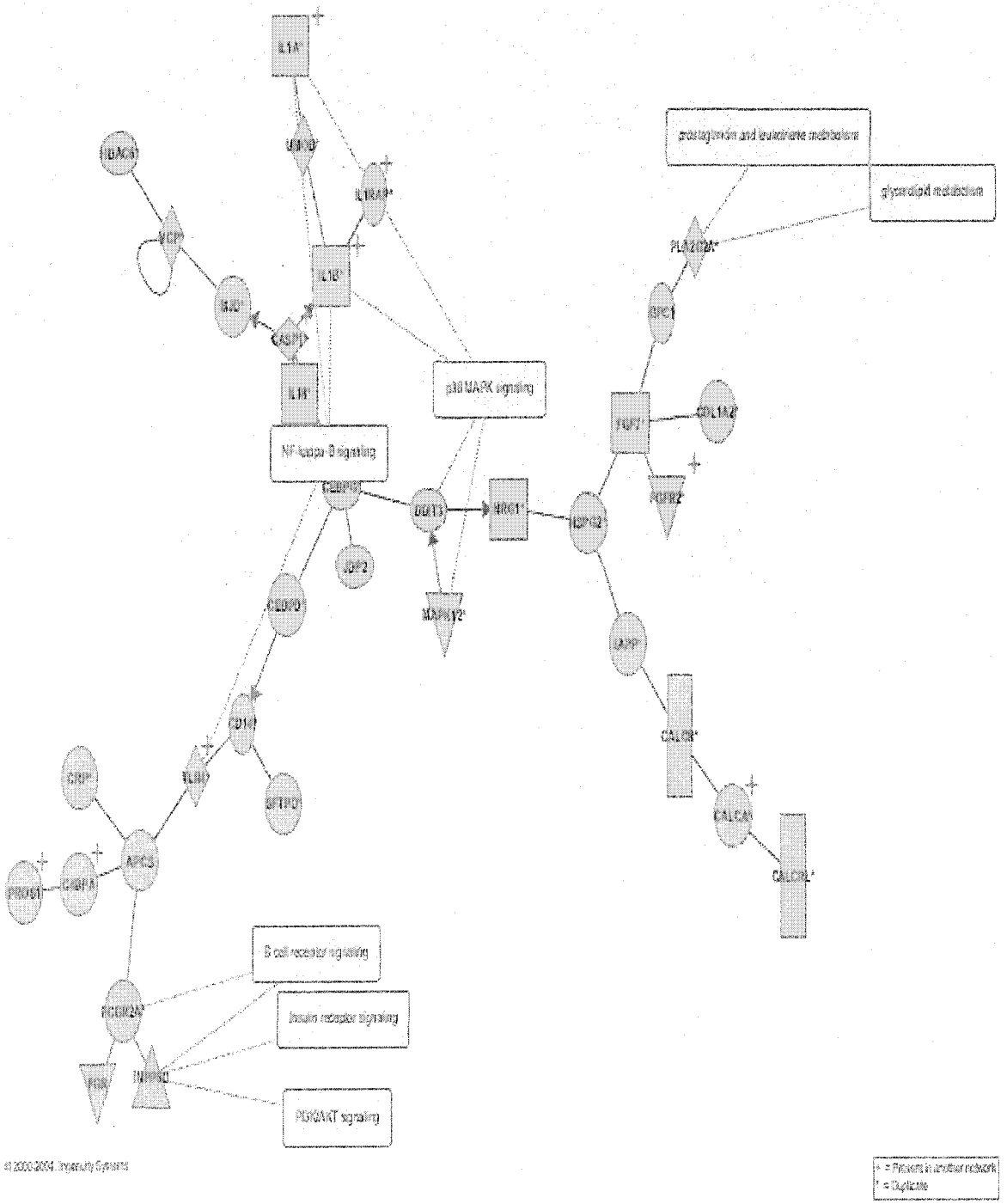
D1

Molecular Signal Transduction Pathways Regulated by BaP



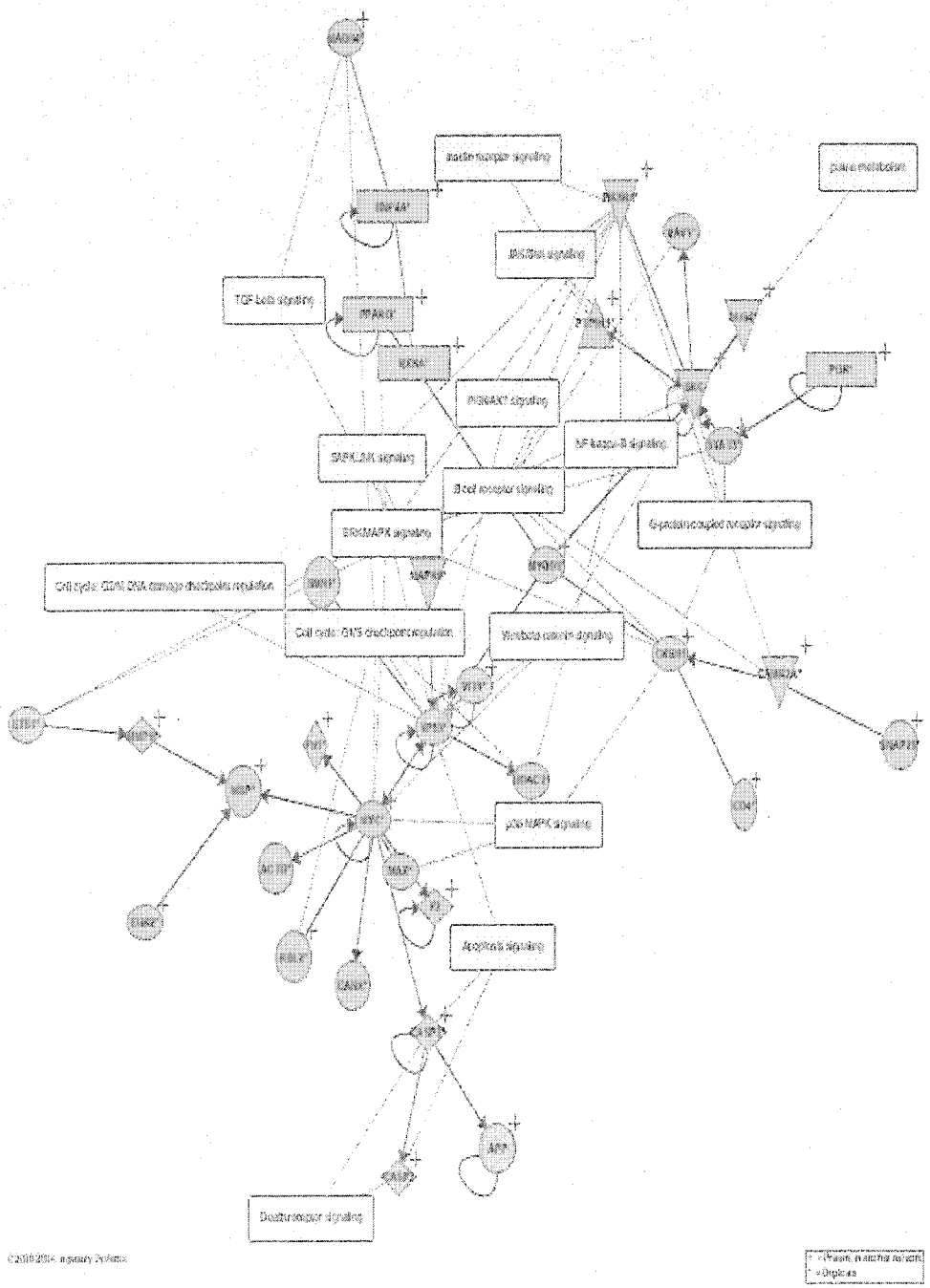
D2

BaP Regulation of *IL-1β* Through the p38 MAPK and NF-kappa B Signaling



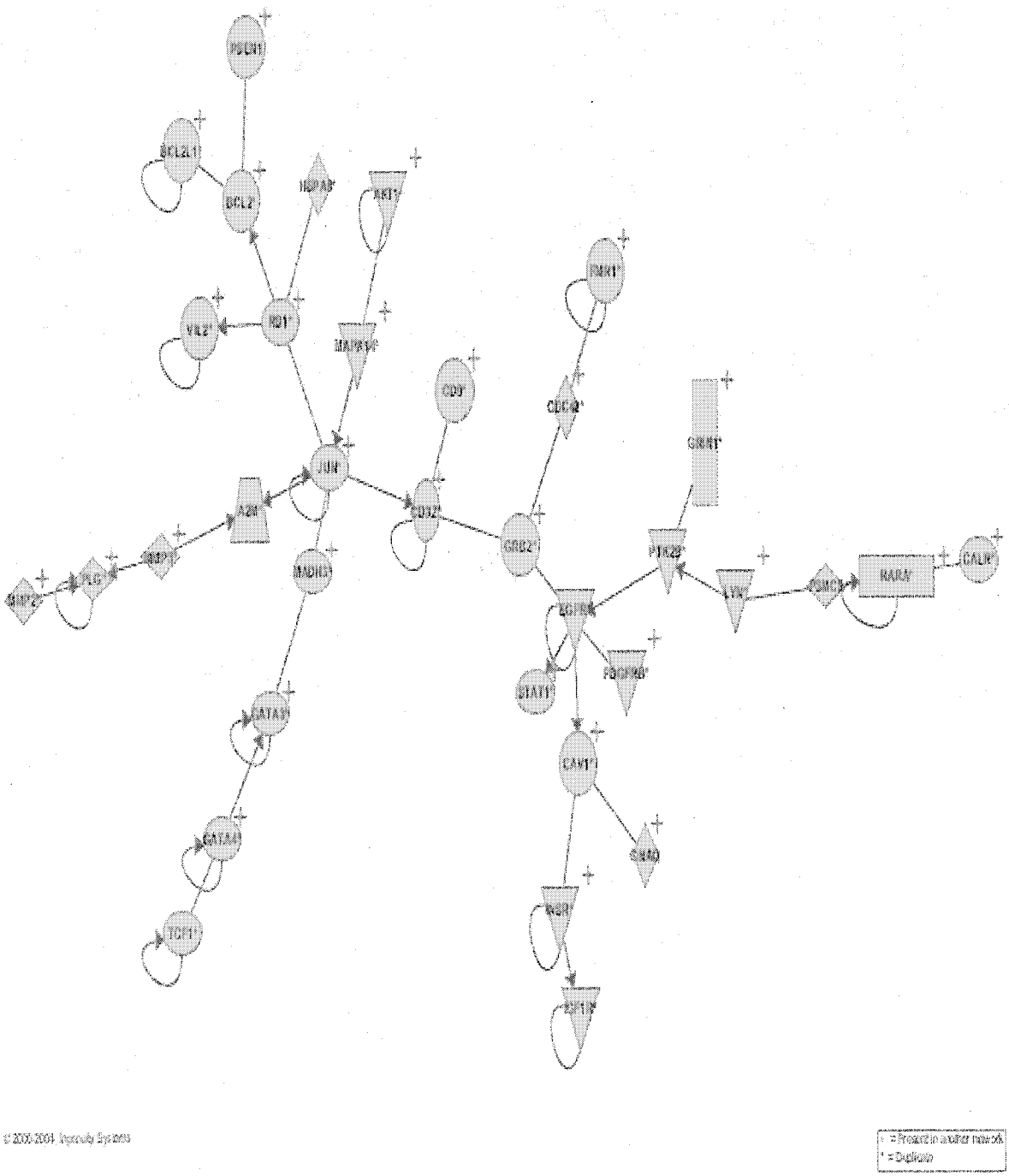
D3

BaP Regulation of *p53* and Other Oncogene Families



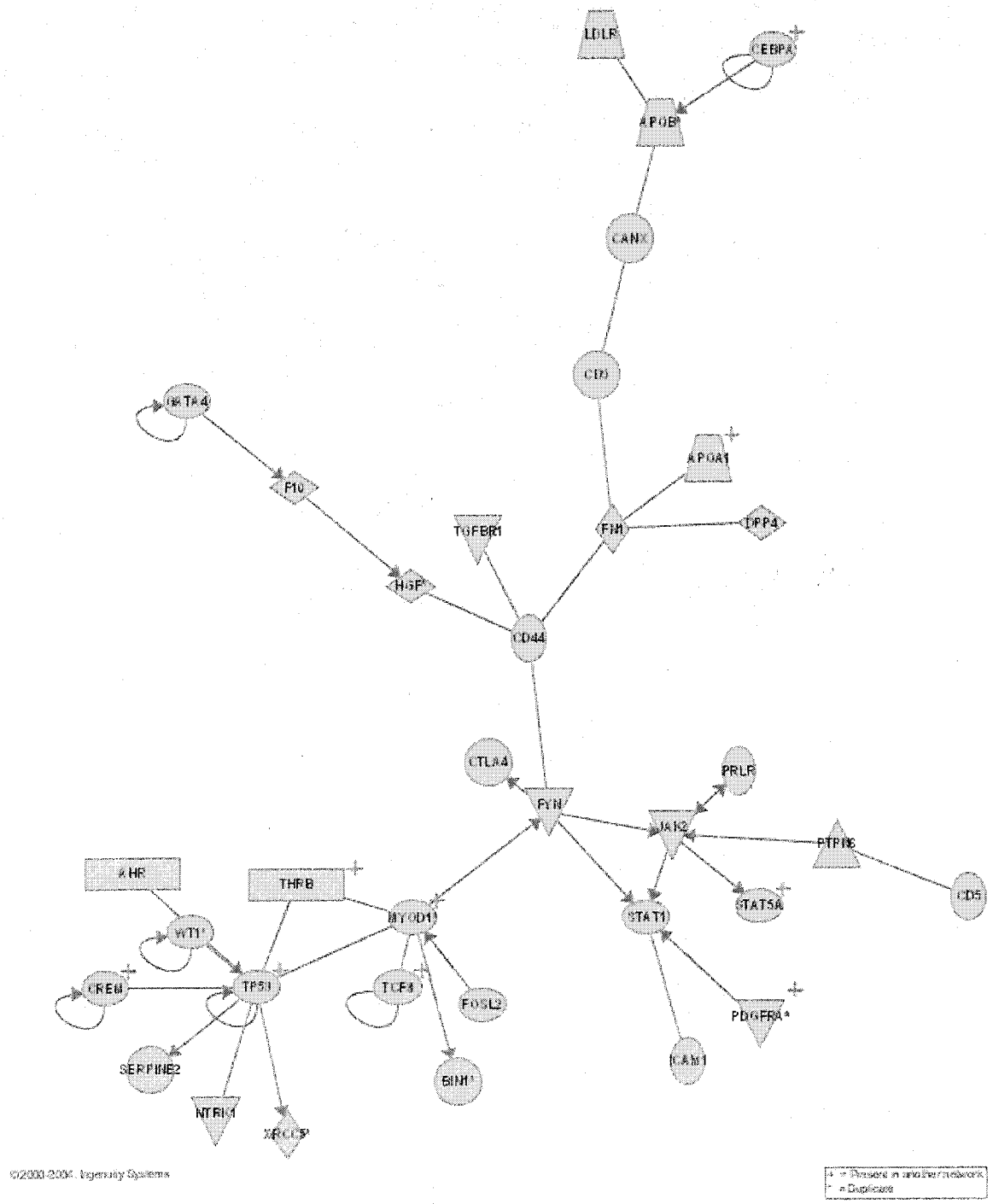
D4

BaP Regulation of Bcl-2 and Other Anti-apoptosis Genes



D5

Ah Receptor Involvement in the Regulation of *p53* and Other Genes



BIBLIOGRAPHY

Afshari, C.A., Nuwaysir, E.F., Barrett, J.C., 1999, Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation, *Cancer Research,* 59: 4759-60

Agency for Toxic Substance and Disease Registry, 1990, *Public health Statement: Benzo (a) pyrene.* Atlanta, GA: Division of Toxicology.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D., 1994, Molecular Biology of the Cell, Third Edition, Garland Publishing, New York, NY

Amin, R.P, Vickers, A.E, Sistare, F, Thompson, K.L, et al, 2004, Identification of putative gene-based markers of renal toxicity, *Environmental Health Perspectives,* 112, 4, 465-79

Arif, J.M and Gupta, R.C., 1996, Detection of DNA-reactive metabolites in serum and their tissue distribution in mice exposed to multiple doses of carcinogen mixtures: role in human biomonitoring, *Carcinogenesis,* 17, 2213 - 19.

Baker, V.A, Harries, H.M, Waring, J.F., Duggan, C.M., Ni, H.A et al., 2004, Clofibrate-induced gene expression changes in rat liver: A cross-laboratory analysis using membrane cDNA arrays, *Environmental Health Perspectives,* 112, 4, 428-38

Barouki, R and Morel, Y., 2001, Repression of Cytochrome P450 1A1 gene expression by oxidative stress: Mechanisms and Biological implications, *Biochemical Pharmacology*, 61, 511-516

Barra, V., 2003, Analysis of gene expression data using functional principal components, *Computer Methods and Programs in Biomedicine* (2004) 75, 1-9

Bartosiewicz, M.J., Jenkins, D, Penn, S., Emery, J., Buckpitt, A., 2001, unique gene expression patterns in liver and kidney associated with exposure to chemical toxicants, *The Journal of Pharmacology and Experimental Therapeutics*, 297, 895-905

Bartosiewicz, M.J., Penn, S., Buckpitt, A., 2001, Applications of gene arrays in environmental toxicology: fingerprints of gene regulation associated with cadmium chloride, Benzo(a)pyrene, and trichloroethylene, *Environmental Health Perspectives*, 109, 1, 71-74

Bartosiewicz, M.J., Trounstine, M., Barker, D., Johnston, R., Buckpitt, A., 2000, Development of a toxicological gene array and quantitative assessment of this technology. *Arch. Biochem Biophys* 376, 66-73

Binkova, B., Giguere, Y., Rossner, P., Dostal, M., Sram, R.J., 2000, The effect of dibenzo[a,1]pyrene and benzo[a]pyrene on human diploid fibroblasts: the induction of DNA adducts, expression of p53 and p21WAF1 proteins and cell cycle distribution, *Mutation Research*, 471, 57-70.

Bjorseth, A., 1983, *Handbook of polycyclic aromatic hydrocarbons.* New York: Marcel Dekker.

Bjorseth, A and Ramdahl, T., 1985, Sources and emission of PAH. In A. Bjorseth and T. Ramdahl (Eds), Handbook of polycyclic aromatic hydrocarbons (Vol.2, pp. 1-20. New York: Marcel Dekker.

Bral, C.M and Ramos, K.S., 1997, Identification of Benzo (a) pyrene inducible *cis*-acting elements within c-Ha-ras transcriptional regulatory sequences, Molecular Pharmacology, 52, 974-982

Brown, K., Buchmann, A., and Balmain, A., 1990, Carcinogen-induced mutations in the mouse c-Ha-ras gene provide evidence of multiple pathways of tumor progression, *Proceedings of the National Academy of Sciences,* USA 87: 538-542

Burczynski, M.E., McMillian, M., Ciervo, J., Li, L., Parker, J.B, and Dunn, R.T., 2000, Toxicogenomics based discrimination of toxic mechanism in HepG2 human hepatoma cells, *Toxicological Science,* 58, 399-415

Burtler, J.P., Post, G.B., Lioy, P.J., Waldman, J.M., Greenberg, A., 1993, Assessment of carcinogenic risk from personal exposure to benzo (a) pyrene in the total human environmental exposure study (THEES). *Journal of the Air and Waste Management Association* 43, 970-977

Burton, G.R., Guan, Y., Nagarajan, R., McGehee Jr, R.E., 2002, Microarray analysis of gene expression during early adipocyte differentiation, *Gene,* 293, 21-31

Canova, S., Degan, P., Peters, L.D., Livingstone, D.R., Voltan, R., Venier, P., 1998,Tissue dose, DNA adducts, oxidative DNA damage and CYP1A-immunopositive proteins in mussels exposed to waterborne benzo[a]pyrene, *Mutation Research 399*, 399, 17-30.

Cavalieri, E.L and Rogan, E.G., 1985, Role of radical cations in aromatic hydrocarbon carcinogenesis, *Environmental Health Perspectives,* 64, 69-84

Chiang, Po-Chang, 2001, Quantification of Benzo (a) pyrene-Guanine Adducts in vitro and *in vivo* tissue samples by LC Tandem Mass Spectrometry, Doctoral Dissertation, Western Michigan University

Clarke, P.A., Poele, R.T, Wooster, R., Workman, P., 2001, Gene expression microarray analysis in cancer biology, pharmacology, and drug development: progress and potential, *Biochemical Pharmacology*, 62, 1311-1336

Cheh, A.M., Chadha, A., Sayer, J.M., Yeh, H.J.C., Yagi, H., Pannell, L.K., and Jerina, D.M., 1993, Structures of covalent nucleosides adducts formed from adenine, guanine, and cytosine bases of DNA and optically active bay-region 3,4-diol 1, 2-epoxides of benz(a)anthracene, *Journal of organic Chemistry,* 58, 4013-4022.

Chen, Y and Tukey, R.H., 1996, Protein Kinase C modulates regulation of the CYP1A1 gene by the aryl hydrocarbon receptor, *Journal of Biological Chemistry,* 271, 42, 26261-26266

Culp, S.J., Gaylor, D.W., Sheldon, W.G., Goldstein, L.S., Beland, F.A., 1998, A comparison of the tumors induced by coal tar and benzo (a) pyrene in a 2-year bioassay. *Carcinogenesis* 19, 117-124.

Dahlquist, K.D., Salomonis, N., Vranizan, K., Lawlor, S.C., Conklin, B.R., *2002* GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways, *Nature Genetics,* 31(1):19-20

Davila, D.R., Davis, D.P., Campbell, K., Cambier, J.C., Zigmond, L.A., and Burchiel, S.W., 1995, role of alterations in Ca^{2+}- associated signaling pathways in the immunotoxicity of polycyclic aromatic hydrocarbons, *Journal of Toxicology and Environmental Health, 45, 101-126*

Denison, M.S and Heath-Pagliuso, S., 1998, The Ah receptor: A regulator of the biochemical and toxicological actions of structurally diverse chemicals, *Bulletin of Environmental Contamination Toxicology, 61, 557-568.*

Edes, T.E., Kwan, S.M., Buckley, C.S., Thornton, W.H., 1992, Tissue vitamin A repletion is impaired by exposure to carcinogen, *International Journal of Cancer*, 50, 99-102

Eisen, M.B, SPELLMAN, P.T, BROWN, P.O and BOTSTEIN, D, 1998, Cluster analysis and display of genome-wide expression patterns, *Proceedings of the National Academy of Sciences USA,* 95, pp. 14863–14868,

Environmental Protection Agency, 1998, U.S. EPA IRIS substance file: Benzo (a) pyrene (BaP). Washington, DC: Author.

Ezendam, J., Staedler, F., Pennings, J., vanderbriel, R.J., Pieters, R., Harleman, J.H., and Vos, J.G., 2004, Toxicogenomics of sub-chronic hexachlorobenzene exposure in brown Norway rats, *Environmental Health Perspectives,* 112, 7, 782-91

Flowers, N.L and Miles, P.R., 1991, Alterations of pulmonary benzo[a]pyrene metabolism by reactive oxygen metabolites, *Toxicology*, 68, 259-274

Goodsaid, F.M., Smith, R.J., Rosenblum, I.Y., 2004, Quantitative PCR deconstruction of discrepancies between results reported by different hybridization platforms, *Environmental Health Perspectives,* 112, 456-64

Grove, D.S, 1999, Quantitative Real-Time Polymerase Chain Reaction for the Core Facility Using TaqMan and the Perkin-Elmer/Applied Biosystems Division 7700 Sequence Detector, *Journal of Biomolecular Techniques,* 10, 11-16

Hall, M.S and Grover, P.L, 1990, Polycyclic aromatic hydrocarbon, metabolism, activation, and tumor initiation, *Chemical carcinogenesis and mutagenesis*, pp 327-372, Springer-Verlag, New York.

Harkin, P.D. 2000, Uncovering functionally relevant signaling pathways using microarray-based expression profiling, *The oncologist*, 5, 501-507

Harvey, R.G., 1991, *polycyclic aromatic hydrocarbons: Chemistry and carcinogenicity.* New York: Cambridge University Press.

Heid, C.A., Stevens, J., Livak, K.J., and Williams, P.M., 1996, Real-Time Quantitative PCR, *Genome Methods,* 6, 986-994

Hossain, M.A., Bouton, C.M., Pevsner, J., Laterra, J., 2000, Induction of vascular endothelial growth factor in human astrocytes by lead. Involvement of a protein kinase C/activator protein-1 complex dependent and hypoxia-inducible factor 1-independent signaling pathway, *Journal of Biological Chemistry,* 169, 105-13

Huang, Q., Dunn, R.T., Jayadev, S., Disorbio, O., Pack, F.D., Farr S.B, et al, 2001, Assessment of cisplatin-induced nephrotoxicty by microarray technology, *Toxicological Science*, 63, 196-207

IARC, 1983, IARC Monographs on the evaluation of carcinogenic risk of the chemicals to humans, Vol. 32, Polynuclear Aromatic Compounds, Part 1: Chemical, Environmental and Experimental Data. Pp. 46-48. 211-373. International Agency for research on cancer, Lyon.

Ingenuity Pathway Analysis, 2004, https://analysis.ingenuity.com/, Ingenuity Pathway Analysis, Ingenuity Inc, CA

Jefcoate, C.R., 1983, Integration of xenobiotic metabolism in carcinogen activation and detoxication. In Biological Basis of Detoxication, J. Caldwell and W.B. Jakoby, eds., pp.31-76, Academic Press, New York

Jeffy, B.D., Chen, E.J., Gudas, J.M., Romagnolo, D.F., 2002, Activation of the aromatic hydrocarbon receptor pathway is not sufficient for transcriptional repression of BRCA-1: requirements for metabolism of benzo[a]pyrene to 7r, 8t-dihydroxy-9t, 10-epoxy-7, 8, 9, 10-tetrahydrobenzo[a]pyrene.

Jeffy, B.D., Chen, E.J., Gudas, J.M., Romagnolo, D.F., 2000, Disruption of cell cycle kinetics by benzo[a]pyrene: Inverse expression patterns of BRCA-1 and p53 in MCF-7 cells arrested in S and G2

Josephy, P.D., 1997, *Molecular Toxicology,* New York: Oxford University Press

Kazerouni, N., Sinha, R., Hsu, C., Greenberg, A., Rothman, N., 2001, Analysis of 200 food items for Benzo(a)pyrene and estimation of its intake in an epidemiologic study, *Food and Chemical Toxicology*, 39, pp 423-436

Kim, H.S., Kwack, S.J., Lee, B.M., 2000,Lipid peroxidation, antioxidant enzymes, and benzo[a]pyrene-quinones in the blood of rats treated with benzo[a]pyrene, *Chemico-Biological Interactions*, 127, 139-150

Kim, K.B and Lee, B.M., 1997,Oxidative Stress to DNA, protein, and antioxidant enzymes (Superoxide dismutase and catalase) in rats treated with benzo(a)pyrene, Cancer letters, 113, 205-212

Klaassen, C., 1995, The basic science of Poisons, *Casarett & Doul's Toxicology,* Fifth edition, Mcgraw Hill Companies.

Ladics, G.S, Kawabata, T.T., Munson, A.E., White, K.L, 1992, Evaluation of Murine Splenic cell type metabolism of Benzo(a)pyrene and functionality in vitro following repeated *in vivo* exposure to Benzo(a)pyrene, *Toxicology and Applied Pharmacology*, 116, 258-266.

Lipshultz, R.J, Fodor, S.P, Gingeras, T.R., and Lockhart, D.J. 1999, High density synthetic oligonucleotide arrays. Nature Genetics 21:20-24

Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T.,Gallo, M.V., Chee, M.S., Mittmann, M, Wang C, Kobayashi M, Horton, H., Brown, E.L., 1996, Expression monitoring by hybridization to high density oligonucleotide arrays, *Nature Biotechnology*, 14, 1675-80

Lodovici, M., Dolara, P., Casalini, C., Ciappellano, S., Testolin, G., 1995, Polycyclic aromatic hydrocarbon contamination in the Italian diet. *Food additives and Contaminant* 12, 703-713

Lu T, Liu J, LeCluyse, E. L, Zhou, Y. S, Cheng,M.L, Waalkes, M.P., 2001, Application of cDNA microarray to the study of arsenic-induced liver diseases in the population of Guizhou, China, *Toxicological Science,* 59, 1, 185-92

Marshall, M.S., 1995, ras target proteins in eukaryotic cells, *FASEB Journal, 9, 1311-1318*

McDonald, D.G and Wood, C.M., 1993, Branchial mechanisms of acclimation to metals in freshwater fish. In: Rankin JC, Jensen, FB, editors. *Fish ecophysiology.* London: Chapman and Hall, p 297-321

Miller, E.C., 1951, Studies on the formation of protein-bound derivatives of 3, 4-benzypyrene in epidermal reactions of mouse skin, *Cancer Research,* 11, 100-108.

Miller, E.C. and Miller, J.A, 1976, the metabolism of chemical carcinogenesis to reactive electrophiles and their possible mechanisms of action in

carcinogenesis. In chemical carcinogens, C.E, Searle, ed., pp. 737-762, American Chemical Society, Washington, DC

Mounho, B.J., Davila, D.R., Burchiel, S.W., 1997, Characterization of intracellular calcium responses produced by polycyclic aromatic hydrocarbons in surface marker-defined human peripheral blood mononuclear cells, *Toxicology and Applied Pharmacology*, 145, 323-330

Mounho, B.J., Davila, D.R., Burchiel, S.W., 1997, Alterations in human B cell calcium homeostasis by polycyclic aromatic hydrocarbons: possible associations with cytochrome p450 metabolism and increased protein tyrosine phosphorylation, *Toxicology and Applied Pharmacology, 149, 80-89.*

Mutch, D.M., Berger, A., Mansourian, R., Rytz, A., Roberts, MA, 2001, The Limit Fold change Model: a practical approach for selecting differentially expressed genes, *BMC Bioinformatics,* 3, 17.

Neumann, N.F and Galvez, F., 2002, DNA microarrays and toxicogenomics: applications for ecotoxicology, *Biotechnology Advances*, 20, 391-419

Ochs, M.F and Godwin, A.K., 2003, Microarrays in Cancer: Research and applications, *Biotechniques supplement,* march 2003.

Overbergh, L., Giulietti, A., Valckx,D, Decallonne, B., Bouillon, R and Mathieu, C., 2003, The Use of Real-Time Reverse Transcriptase PCR for the Quantification of Cytokine Gene Expression, *Journal of Biomolecular Techniques,* 14, 33-43

Pap, E and Csaba, G, 1994, Benzypyrene treatment in adulthood increases the testosterone level in neonatally steroid (Allyestrenol)-treated male rats, *General Pharmacology*, Vol. 25, No.8, pp. 1699-1701.

Pappas, P., Sotiropoulou, M., Karamanakos, P., Kostoula, A., Levidiotou, S., Marselos, M., 2003, Acute phase response to benzo[a]pyrene and induction of rat ALDH3A1, *Chemico-Biological Interactions, 143-144, 55-62*

Parrish, A.R., Fisher, R, Bral, C.M., Burghardt, R.C., Gandolfi, A, J., Brendel, K., Ramos, K.S., 1998, Benzo(a)Pyrene-Induced Aterations In Growth-Related gene Expression and Signaling in Precision-cut adult rat liver and kidney slices, *Toxicology and Applied Pharmacology*, 152, 302-308.

Pei, X., Nakanishi,Y., Inoue, H., Takayama, K., Bai, F., Hara, N., 2002, Polycyclic aromatic hydrocarbons induce IL-8 expression through nuclear factor kB activation in A549 cell line, cytokine, 19, 5, 236-241.

Pessah, I.N., Beltzner, C., Burchiel, S.W., Sridhar, G., Penning, T., Feng, W., 2001, A bioactive metabolite of benzo[a]pyrene-7,8-dione, selectively alters microsomal Ca^{2+} transport and ryanodine receptor function, *Molecular Pharmacology,* 59, 3, 506-513.

Ramet, M., Castren, K., Jarvinen, K., Pekkala, K., Turpeeniemi-Hujanen, T., Soini, Y., Paakko, P and vahakangas, K., 1995, p53 protein expression is correlated with benzo[a]pyrene-DNA adducts in carcinoma cell lines, *Carcinogenesis*, vol.16, 9, 2117-2124

Ramesh, A., Inyang, F., Hood, D.B., Knuckles, M.E., 1999, Aryl hydrocarbon hydroxylase activity in F-344 rats subchronically exposed to benzo(a)pyrene and fluoranthene through diet, *Journal of Biochemical Molecular Toxicology,*14, 3, 155-161.

Rao, V.L.R., Bowen, K., Dhodda, V.K., Song, G., Franklin, J.L., Gavva, N.R., Dempsey, R.J., 2002, gene expression analysis of spontaneously hypertensive rat cerebral cortex following transient focal cerebral ischemia, *Journal of Neurocehmistry*, 83, 1072-1086

Rininger, J.A., DiPippo, V.A., Gould-Rothberg, B.E, 2000, Differrential gene expression technologies for identifying surrogate markers of drug efficacy and toxicity, *Drug Discovery Today,* Vol. 5, 12, 560-568

Rogan, E.G., Ramakrishna, N.V.S., Higginbotham, S., Cavalieri, E.L., Jong, H., Jankowiak, R., and Small, G.J., 1990, Identification and quantitation of 7-(benzo(a)pyrene-6-yl)-guanine in the urine and feces of rats treated with benzo(a)pyrene, *Chemical Research Toxicology,* 3, 441-444

Salas, V.M and Burchiel, S.W, 1998, Apoptosis in Daudi human B cells in response to benzo[a]pyrene and benzo[a]pyrene-7,8-dihydrodiol, *Toxicology and Applied Pharmacology*, 151, 367-376

Schena, M., 2000, Microarray biochip technology, Westborough, MA: *Biotechniques Press*

Schena, M., Shalon, D., Davis, R., Brown, P., 1995., Quantitative monitoring of gene expression patterns with a complemetary DNA microarray, *Science* 270: 467-470.

Schmidt, J.V and Bradfield, C.A., 1996, Ah receptor signaling pathways, *Annual Review of Cell Development Biology*, 12, 55-89

Shirasawa S., Furuse, M., Yokoyama, N., and Sasazuli, T., 1993, altered growth of human colon cancer cell lines disrupted at activated Ki-ras, *Science, 260, 85-88*

Simmonds, M.A. 2003, Cancer statistics, 2003: Further decrease in mortality rate, increase in persons living with cancer. CA Cancer J.Clin. 53:4

Szczekilik, A., Szczeklik, J., Galuszka, Z., Musial, J., Kolarzyk, E, and Targosz, D,1994, Humoral immunosuppression in men exposed to polycyclic aromatic hydrocarbons and related carcinogens in polluted environments. *Environmental Health Perspectives* 102: 302-304

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., Golub, T.R., 1999, interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoitic differentiation, *Proceedings of the National Academy of Sciences*, 96, 2907-2912.

Thomas, R. P., Guigneaux, M., Wood, T., Evers, M., 2002, Age –Associated changes in gene expression patterns in the liver, *Journal of Gastrointestinal Surgery*, 6, 445-454

Tian, Y., Ke, S., Denison, M.S., Rabson, A.B., Gallo, M.A., 1999, Ah Receptor and NF-kB Interactions, a potential mechanism for Dioxin toxicity, *Journal of Biological Chemistry*, vol. 274, 1, pp 510-515

Toronen, P., Kolehmainen, M., Wong, G., Castren, E., 1999, Analysis of gene Expression data using self-organizing Maps, *FEBS Letters* (451), 142-146

Trombino, A.F., Near, R.I., Matulka, R.A., Yang, S., Hafer, L.J., Toselli, P.A., Kim, D.W., Rogers, A.E., Sonenshein, G.E, Sherr, D.H., Expression of the aryl hydrocarbon receptor/transcription factor(AhR) and AhR-regulated CYP1 gene transcripts in a rat model of mammary tumorigenesis.

Vinggaard, A.M., Hnida, C., Larsen, J.C., 2000, Environmental polycyclic aromatic hydrocarbons affect androgen receptor activation in vitro, *Toxicology,* 145, 173-183

Waalkes, M.P and Ward, J.M.,1994, Carcinogenesis: Target Organ Toxicology Series, New York: Raven Press.

Ward, M.H., Sinha, R., Heineman, E.F., Rothman, N., Markin, R., Weisenburger, D.D., Correa, P., Zahm, S.H., 1997, Risk of adenocarcinoma of the stomach and esophagus with meat cooking method and doneness preference, *International Journal of Cancer*, 71, 14-19.

Waring, J. F., Jolly, R.A., Ciurlionis, R., Lum, P.Y., Praestgaard, J.T., Morfitt, D.C., Buratto, B., Roberts, C., Schadt, E., Ulrich, R.G., 2001, Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles, Toxicology and Applied Pharmacology, Vol. 175, 28-42

Wei, C., Caccavale, R.J., Weyand, E.H., Chen, S., Iba, M.M., 2002, Induction of CYP1A1 and CYP1A2 expressions by prototypic and atypical inducers in the human lung, *Cancer Letters,* 178, 25-36.

Weyand, E.H., Chen, Y.C., Wu, Y., Koganti, A., Dunsford, H.A, Rodriquez, L.V., 1995, Differences in the tumorigenic activity of a pure hydrocarbon and a complex mixture following ingestion: benzo (a) pyrene vs. manufactured gas plant residue. *Chemical Research in Toxicology 8, 949-954*

Yan, Z., Subbaramaiah, K., Camilli, T., Zhang, F., Tanabe, T., McCaffrey, T.A., Dannenberg, A.J., Weksler, B.B., 2000, BaP induces the transcription of cyclooxygenase-2 in vascular smooth muscle cells, *Journal of Biological Chemistry*, Vol. 275, 7, pp 4949-4955

Yeung, K.Y and Ruzzo, W.L., 2001, Principal component analysis for clustering gene expression data, *Bioinformatics*, 17, 9, 763-774

Yoneda, K., Peck, K., Chang, M.M., Chmiel, K, Sher, Y.P., Chen, J. et al., 2001, Development of high-density DNA microarray membrane for profiling smoke-and hydrogen peroxide-induced genes in a human bronchial epithelial cell line, *American Journal of Respiratory Critical care medicine,* 164, S85-9