uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Natalie Ann Osika
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biology)
GRADE / DEGREE

Department of Biology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Identification of Candidate Exposure Biomarkers for Toxicants in Complex Environmental Matrices

TITRE DE LA THÈSE / TITLE OF THESIS

Paul White
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Marc Ekker

Francis Pick

Carole Yauk

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Identification of Candidate Exposure Biomarkers for Toxicants in Complex Environmental Matrices

Natalie Ann Osika, B.Sc.

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the M.Sc. degree in Chemical and Environmental Toxicology

Department of Biology
Faculty of Science
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-79692-4
Our file Notre référence
ISBN: 978-0-494-79692-4

Abstract

The term "biomarker" refers to an indicator employed to monitor the interactions between a xenobiotic and an individual, provide information of an individual's biologic state, or refer to an individual's susceptibility to a xenobiotic, condition or disease. They are currently employed to assess exposure to, or effects from, environmental toxicants. It is important to have biomarkers that can accurately and precisely assess complex mixture exposures, since environmental exposures to single substances rarely occur outside the laboratory setting, and complex exposures are often associated with specific occupational or environmental settings. This project employed microarray technology and subsequent bioinformatic analyses to identify candidate biomarkers of *in vitro* exposure to coal tar extract. Murine pulmonary epithelial cells were exposed to coal tar extract and gene expression was assessed using microarrays. The results showed that expression of genes involved in steroidogenesis, cytokine-cytokine interaction, angiogenesis, as well as several genes that code for secreted proteins, were altered by coal tar exposure. The entire work constitutes the initial stages of a project that will ultimately identify gene expression and secretome profiles that are specific to coal tar exposure.

Résumé

L'expression « marqueur biologique » fait référence à un indicateur utilisé pour surveiller les interactions entre un xénobiotique et un individu, fournir l'information relative à l'état biologique d'un individu ou faire référence à la vulnérabilité d'un individu à une maladie xénobiotique. Le marqueur biologique est utilisé pour évaluer l'exposition aux agents toxiques environnementaux ou leurs effets. Il est important d'avoir des marqueurs biologiques pouvant évaluer avec précision les expositions aux mélanges complexes puisque les expositions environnementales à une substance unique se produisent rarement à l'extérieur du laboratoire et les expositions complexes sont souvent associées à des contextes professionnels ou environnementaux. Dans le cadre de ce projet, la technologie du microréseau a été utilisée et des analyses bioinformatiques subséquentes ont été effectuées pour cerner les marqueurs biologiques candidats d'une exposition *in vitro* à un extrait de goudron de houille. Des cellules épithéliales pulmonaires murines ont été exposées à l'extrait de goudron de houille et l'expression génétique a été évaluée au moyen de microréseaux. Les résultats ont montré que l'expression des gènes impliqués dans la stéroïdogénèse, l'interaction cytokine-cytokine, l'angiogénèse ainsi que plusieurs gènes produisant le code des protéines sécrétées ont été altérés en raison d'une exposition au goudron de houille. Tout le travail constitue les premières étapes d'un projet qui permettra de cerner l'expression des gènes et d'établir les profils sécrétomiques propres à l'exposition au goudron de houille.

Acknowledgements

There were many people who supported and contributed to the research within this thesis. Firstly I would like to thank Dr. Paul White for taking me on as his Master's student and allowing me to work on such an interesting and challenging project. This was a tremendous privilege that has created many opportunities and success in my career that I have been fortunate enough to be able to enjoy, even before graduation.

I would like to express my sincere gratitude to my lab partner and mentor, Dr. Mamady Hapsatou, who contributed greatly to the design of the project and the toxicity assays.

I would also like to express my gratitude to Dr. Carole Yauk, for providing so much guidance and sharing her expertise in "omics" research.

I would like to thank everyone at the Mechanistic Studies Division, especially Lynn Berndt-Weis, Andrea Rowan-Carroll and Dr. Sabina Halappanavar for taking their time to teach me all computer programs and microarray protocols.

I am grateful to Andrew Williams and Morie Malowany who completed all the statistical analysis for my microarray experiment, and were kind enough to patiently answer my exceedingly numerous questions.

I would like to thank my Mom and Dad who have always provided me much encouragement and support throughout my education. Last but not least I would like to thank Bill and Eddy for getting me through the tough times.

Funding was provided by the Government of Canada's Genomics Research and Development Initiative and the Canadian Research Strategy for Biotechnology.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BaP- Benzo[*a*]pyrene

CYP- Cytochrome P450

DMSO- Dimethyl sulfoxide

FDR- False Discovery Rate

GO- Gene ontology

HACs- heterocyclic aromatic compounds

KEGG- Kyoto Encyclopedia of Genes and Genomes

LDH- Lactate dehydrogenase

NIST-National Institute of Standards and Technology

NPRI- National Pollutant Release Inventory

PAH- Polycyclic Aromatic Hydrocarbon

PBS- phosphate buffered saline

PCBs- Polychlorinated biphenyls

PSA- prostate specific antigen

XTT- sodium 3'-[1-phenylaminocarbonyl]-3, 4-tetrazorium)-bis [4-methoxy-6-nitro] benzene sulfonic acidhydrate

Chapter 1: Introduction

1 Introduction

1.1 Biomarkers

The term "biomarker" refers to an indicator employed to monitor the interactions between a xenobiotic and an individual, provide information of an individual's biologic state, or refer to an individual's susceptibility to a xenobiotic, condition or disease. A good biomarker should be a precise indicator of an exposure, effect or disease; as well as be sensitive, non-invasive, cost-effective and have a long half life with the ability to show cumulative exposure (Dunn 2008).

Biomarkers of susceptibility are genetic or non-genetic factors (e.g., environment or diet) that alter an individual's risk of disease or xenobiotic effects (Conti 2008). The ability to metabolize xenobiotics, which is genetically determined by polymorphisms, is a popular topic in the environmental toxicology literature (Conti 2008, Duffus 2006). For example, a polymorphism of the gene N-acetyltransferase (*Nat2*) permits more rapid acetylation of aromatic amines, lowering the risk of amine-induced bladder cancer (Duffus 2006). Another family of genetic polymorphisms responsible for altered susceptibilities are those that code for different isoforms of the Cytochrome P450 (*Cyp*) enzymes that are responsible for metabolizing many endogenous and exogenous aromatic compounds (Sigel 2007). The aryl hydrocarbon receptor (AhR) controls the expression of many of the *Cyp* genes, including *Cyp1a1*, *Cyp1a2*, and *Cyp1b1*. Substances such as polycyclic aromatic hydrocarbons (PAHs) can serve as AhR agonists, and may increase *Cyp* expression and activity leading to the formation of harmful PAH metabolites that can increase the risk of cancer in susceptible individuals (Hong & Yang 1997).

Biomarkers of exposure are important in biomonitoring because they reflect bioavailability and the magnitude of the internal dose of a xenobiotic. They can be used to monitor humans in occupational settings, and although they are not used to measure the adverse health effects caused by an exposure, they can be used to predict the likelihood of disease (Henderson 1995). Biomarkers of exposure can be a measure of the xenobiotic itself, or a metabolite of a xenobiotic, in bodily fluids such as the blood or urine. Examples include chromium in urine as an indicator of exposure to chromium sulphate in leather tanning (Aitio *et al.* 1984), trichloroethanol in blood as an indicator of trichloroethylene exposure cotinine, a metabolite of nicotine, in urine as an indication of nicotine exposure (Heinrich-Ramm *et al.* 2002), and 1-hydroxypyrene in urine, as an indicator of PAH (i.e., pyrene) exposure (Jongeneelen *et al.* 1988).

A biomarker of effect is a biochemical change, or early reversible effect, that can be used as an indicator of exposure or adverse health effect. They are frequently used as preclinical markers of pathology elicited by an exposure (Marek & Wocka-Marek 1994). Examples of biomarkers of effect include, N-acetylglucosaminidase in urine as a biomarker of mercury exposure/effect (Marek & Wocka-Marek 1994), and red blood cell acetylcholinesterase activity as a marker of organophosphate exposure/effect (Thiermann 2005). Biomarkers of effect can sometimes be used as prognostic indicators of disease. For example, the presence of B2-microglobulin in the urine denotes damage to kidney tubules from cadmium exposure (Karakaya *et al.* 1993). This example is very specific to the type of exposure that elicits the adverse effect; however, there are many biomarkers of effect that are not specific to a single type of exposure. For example, the frequency of

micronuclei in reticulocytes denotes exposure to any genotoxic chemical that can cause chromosome damage (Fenech 1998).

Biomarkers of effect have frequently been used in chronic disease epidemiology, and clinical biomarkers are often useful for early detection and diagnosis of cancer. Examples include prostate-specific antigen (PSA), used in the diagnosis of prostate cancer (Fleshner & Lawrentschuk 2009), mammaglobin-2 (MGB2), which can detect metastasis of breast and ovarian cancer (Adib *et al.* 2004), and α-fetoprotein, which is used for the detection of hepatocellular carcinomas (Abelev 1971). PSA is an androgen-regulated serine protease that is produced by both prostate epithelial cells and prostate carcinomas (Balk *et al.* 2003). It exists in small quantities in all men, but an elevated serum level is currently the best early indicator of prostate cancer (Andriole *et al.* 2009). MGB2 is a protein expressed by the adult mammary gland. The presence of this protein in the lymph nodes has been shown to be strongly correlated with micrometastases of breast carcinomas (Aihara 2004), although research and validation to determine its utility as a clinical biomarker is ongoing. α-fetoprotein is normally produced during gestation, and production declines with age; however, it is present in the serum of children with hepatoblastoma and malignant germ cell tumors (Wu *et al.* 1981, Wilkins 2006). Although it is used as a tool to for early detection of cancer, it is more useful for monitoring the efficacy of cancer treatments (Wu *et al.* 1981).

It is important to note that there is overlap between biomarkers of exposure and biomarkers of effect. For example, if a pathological biochemical change is caused by an exposure, and the purpose of the biomarker is to identify the exposure, then the biomarker of effect may be referred to as a biomarker of exposure. This is quite common for

biomarkers of environmental exposures. For example, plasma vitellogenin concentration in male teleost fish is indicative of a pathophysiologic change induced by exposure to environmental estrogens (Heppell *et al.* 1995). Another example of a biomarker of exposure and effect is urinary AFB-N-7-guanine adduct, an Aflatoxin B_1 nucleoside adduct, which is indicative of both Aflatoxin B_1 exposure and DNA damage (Groopman *et al.* 1992).

The research described in this thesis contributes to the discovery of biomarkers of exposure and/or effect related to an environmental exposure. Existing biomarkers of effect for environmental exposures include altered gene expression of cytochrome *p4501a1* and *1b1*, increased micronucleus frequency in target tissues, urinary 1-hydroxypyrene from PAH exposures, and urinary cotinine from nicotine exposure. These biomarkers, with the exception of cotinine, are generally not specific enough to determine the exact source and nature of the exposure. Moreover, many biomarkers are indicators of exposure to a chemical or chemical family that originates from multiple sources, making the precise origin of the exposure impossible to identify. For example, 1-hydroxypyrene indicates exposure to pyrene (Roggi *et al.* 1997), originating from many possible sources including vehicle exhaust, coal tar, coke plant emissions, wood smoke, and urban air (Castano-Vinyals *et al.* 2004). Since it is a metabolite of nicotine, cotinine is one of the few biomarkers that is specific to a single complex environmental exposure (i.e.tobbaco smoke) (Florescu *et al.* 2009, Benowitz & Jacob 1994). The cancer risk from environmental exposure to tobacco is noteworthy, and cotinine in body fluids has been extensively evaluated and employed as a biomarker of exposure (Benowitz 1996).

It is important to have biomarkers that can accurately identify complex mixture exposures, since exposures to single substances rarely occur outside the laboratory setting, and complex exposures are often associated with specific occupational or environmental settings. Efforts to establish correlations between occupational exposures to a specific PAH source, and an increased risk of adverse effect risk, such as cancer, have been largely unsuccessful. For example, coke plant workers have been shown to have an increased cancer risk from chronic PAH exposure (Costantino *et al.* 1995), but the traditional biomarkers of exposure to PAHs, including 1-hydrodroxypyrene in urine, PAH-albumin adducts in serum and CYP expression, are not specific enough to separate occupational exposures from other PAH sources (e.g., tobacco smoke and air pollution) (Pan *et al.* 1998). A study that examined non-smoking Danish bus drivers and postal workers observed significantly elevated PAH-albumin adducts in a suburban group of bus drivers and postal workers, in comparison with groups in the central part of Copenhagen, Denmark (Autrup *et al.* 1999). This unexpected result (i.e., elevated PAH albumin adducts in subjects from suburban areas) has been seen in other studies (Autrup *et al.* 1995, Nielsen *et al.* 1996), and is thought to be related to exposures to emissions from the combustion of wood and cereal straws used for heating. A biomarker or collection of biomarkers, that is specific to an occupational exposure would be useful for monitoring occupational health and safety, and for litigation on behalf of negatively affected workers. Likewise, the ability to use a specific biomarker to assign a disease to a specific exposure event would be useful for government regulators endeavouring to safeguard public health.

The development of genomic technologies such as global gene expression profiling has tremendously enhanced biomarker research. For example, DNA microarray technology

has been used in clinical biomarker research to identify multi-gene signatures that accurately predict metastatic tumor growth (Goncalves *et al.* 2006). The recurrence of breast cancer can be expected in only 20-30% of young women with early stage breast cancer, however, 85-90% of breast cancer patients receive aggressive cytotoxic adjuvant chemotherapy (van't Veer & Bernards 2008). Therefore, a large proportion of women receiving aggressive cancer chemotherapy will not benefit from the treatment and suffer side effects. Consequently, researchers at Rosetta Inpharmatics used global gene expression profiling to identify and develop clinically-relevant prognostic biomarkers of tumor development and malignancy. The analyses of van t'Veer et al. established a poor prognosis gene expression signature consisting of multiple genes that regulate cell cycle, invasion, metastasis and angiogenesis (van 't Veer *et al.* 2002). Their DNA microarray analysis of 98 tumors, including 34 from patients whose cancer metastasized within 5 years, and 44 from patients who remained disease free after 5 years, employed a hierarchical clustering algorithm of five thousand significant genes to separate tumours into prognostic groups (i.e., poor prognosis and good prognosis) on the basis of genomic similarity (van 't Veer *et al.* 2002). The work ultimately led to the development of specialized microarrays employed to derive the Mamma Print™, an expression profile of 70 genes that forms the basis for categorization of tumour samples into prognostic groups (i.e., high or low risk of recurrence following surgical removal)(Agendia, 2009).

Toxicogenomics has employed gene expression data to determine how the genome responds to chemical toxicants and other environmental stressors, thus providing an improved understanding of the role of genes, and gene-environment interactions, in determining the risk of adverse effects and disease (Hamadeh *et al.* 2002a). Given the vast

number of chemical toxicants that exist, the collective expression of the genes corresponding to a specific type of compound is an effective strategy for discovering useful, precise biomarkers of exposure and pathophysiologic change. Indeed, toxicogenomics, and the use of DNA microarrays for gene expression profiling, has already led to the discovery of multi-gene markers of exposure to specific types of toxicants. For example, Hamadeh et al. (2002) determined that agents from the same chemical class (e.g., peroxisome proliferators) could produce similar patterns of gene expression in rat liver tissue. The results revealed that peroxisome proliferators induced similar patterns of gene expression, while phenobarbital induced a very different pattern of gene expression (Hamadeh *et al.* 2002b). In addition, analysis of the genomic signature provided important mechanistic information on underlying toxicological mechanisms.

Although the aforementioned research provided valuable information, gene expression profiles in target tissues are of limited use for human exposure monitoring since the samples required (i.e., samples of target tissues) cannot be easily obtained. Although a target tissue approach is readily acceptable for tumour profiling, its use for monitoring of occupational and environmental exposures is clearly not practical. In contrast, proteomic profiling of serum provides a more direct (Mbeunkui *et al.* 2006) and non-invasive alternative for biomonitoring. As a result, the *secretome*, the fraction of the proteome that is secreted by a cell/tissue (Greenbaum *et al.* 2001), is of great interest to biomarker researchers. The secretome is responsible for regulating a variety of biological processes, such as cell signaling, growth and apoptosis (Pellitteri-Hahn *et al.* 2006, Greenbaum et al. 2001), and can be isolated from blood plasma and cerebral spinal fluid (Suk 2010). However, techniques for the routine analysis of proteomic changes elicited by

environmental or occupational exposures are not nearly as mature as genomic techniques, and they are therefore, not as practical for biomarker discovery. Moreover, the complex nature of bodily fluids such as serum confounds routine measurement of secretome changes in these fluids. For example, many elements of the plasma secretome are present in small amounts relative to highly abundant proteins such as albumin and immunoglobin G (Ding *et al.* 2009). These complications make genomic directed analysis of the proteome in general, or the secretome, more appealing (Klee *et al.* 2004).

1.2 DNA MicroarrayTechnology

DNA microarrays were first mentioned in the genome issue of Science in 1995 (Schena *et al.* 1995). Since then they have become a powerful tool for systematic parallel analysis of large numbers of genes. They evolved from Northern blotting, whereby genomic samples of mRNA can be probed for specific genes of interest. The problem with this technique is that only a limited number of genes can be analyzed simultaneously. One of the greatest advantages of microarrays is their ability to permit simultaneous monitoring of gene expression changes for enormous numbers of "arrayed" genes, even an entire genome, in a single experiment. DNA microarrays consist of thousands of arrayed DNA probes, each of which represents a specific gene, fixed in a known location to a glass or silicon slide. The DNA probe on the microarray slide binds to complementary, fluorescently labeled mRNA copies in a given sample, and the intensity of the resulting signal for a given spot (or feature) is indicative of the abundance of the mRNA in the genomic sample (i.e., gene expression). The intensity can be compared to a standard to quantify the amount and direction of changes in gene expression elicited by the treatment

or condition. For toxicogenomics, the study of gene expression changes elicited by exposures to toxic agents, having the entire genome represented on a single chip is useful for the discovery of previously unidentified or unexpected changes in gene expression (i.e., eliminates *a priori* assumptions about the nature of the response) . As such, the technology provides a powerful tool for the identification of novel biomarkers of exposure and effect. Since its introduction, microarray technology has become increasingly popular; however, its potential in environmental health sciences and applied toxicology has not yet been fully realized.

1.3 COAL TAR and Polycyclic Aromatic Hydrocarbons

Polycyclic aromatic hydrocarbons are a group of organic chemical compounds that are lipophilic, stable, and not very volatile (Scott 2002). They are a group of multi-ringed compounds formed during the incomplete combustion of organic matter. PAHs were included in first list of *priority substances* under section 11 of CEPA (Canadian Environmental Protection Act) because the larger, five and six-ring compounds are mutagenic and carcinogenic when metabolized to electrophilic intermediates such as diol epoxides (CEPA 1988).

PAH-induced mutagenesis can initiate carcinogenesis. Mutagenesis occurs when the PAH diol epoxides covalently bind to DNA, forming stable adducts that can lead to permanent changes in sequence (i.e., mutations). Several noteworthy PAHs, which are thought to be responsible for a substantial portion of the carcinogenic activity of PAH mixtures (e.g., benzo[*a*]pyrene, benz[*a*]anthracene, benzo[*b*]fluoranthene), include those that contain a bay- or fjord-regions, and can be metabolized to bay- or fjord-region diol-

epoxides that readily interact with DNA (see Figure 1). The fjord region has been shown to have stronger tumourigenic activity than the bay-region due to a more pronounced planar interference of the other rings in the aromatic compound caused by the comparatively less open, crowded shape of the fjord region (Carrell *et al.* 1994). Moreover, the fjord region diol-epoxide has been shown to induce a type of DNA adduct that is not easily removed by DNA repair pathways (Buterin *et al.* 2000), and unrepaired adducts can cause mutations in critical genes that initiate carcinogenesis (Hecht *et al.* 1994).

The mutagenic effects of PAHs have been known for some time, and skin cancer was reported among workers exposed to PAHs employed in coal tar distilleries as early as 1876 (Bell 1876). Research conducted by Kenneway et al in the 1930s identified benzo[*a*]pyrene (BaP) as a potent carcinogen in coal tar. It exhibits strong mutagenic properties in laboratory animals, and since its identification, has been one of the most extensively studied carcinogens (Phillips 1983).

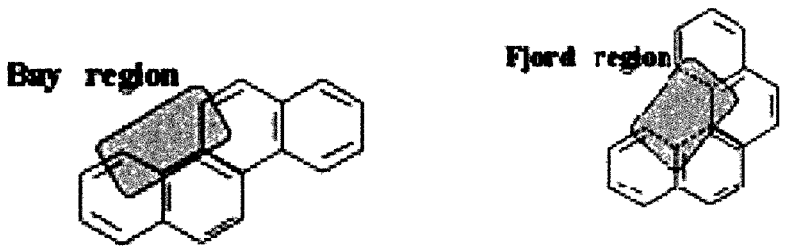


Figure 1.1 The Bay and Fjord regions of a polycyclic aromatic hydrocarbon
Modified from :

http://www.nyu.edu/projects/geacintov/images/WEB_Figures/WEB_PAH1.jpg

Another class of aromatic compounds that is closely related to PAHs, and which may exhibit similar mutagenic effects, is the heterocyclic aromatic compounds (HACs). These compounds feature a nitrogen, sulfur or oxygen atom that replaces one of the carbon atoms in a ring, for example oxazole and tetrazole (CEPA, 1993). Major anthropogenic sources of both PAHs and HACs include a variety of industrial by-products and emissions such as coal tar and coal tar creosote, passive and active tobacco smoke, petroleum and petroleum derivatives, motor vehicle exhaust, domestic heating emissions, and forest fire emissions (Scott 2002). PAHs have been found at levels up to 20 $\mu g/m^3$ in the atmosphere, coal furnace emissions can contain over 1000 $\mu g/m^3$, and cigarette smoke contains about 100 $\mu g/m^3$(Manahan 2005). Non-anthropogenic sources include volcanic eruptions or natural vegetation fires (Krauss *et al.* 2005). According to the National Pollutant Release Inventory NPRI, the total amount of PAHs released from industrial facilities across Canada was 645.63 tonnes in 2008 (NPRI 2008).

Increases in the atmospheric concentration of PAHs are often attributed to the industrial revolution. In a study by Hites et al.(Hites 1980), a coastal sediment core collected in New England was analyzed and total PAHs showed a profile of increasing PAH concentration over time. Moreover, the profile showed declines in industrial production during the Great Depression, increases in production associated with World Wars I and II, the and the introduction of cleaner fuels and environmental regulations that started in the Cold War era.

Atmospheric PAHs are almost always associated with the solid phase, especially adsorbed to soot. Environmental behaviour and bioavailability of PAHs are dependant on

source and route of exposure, and PAH exposures generally occur via inhalation, ingestion or dermal contact. PAHs on smaller inhaled particles can be absorbed through the bronchial or alveolar epithelia, and the rate of absorption for particle-associated PAHs depends on the size of the particles (Yang *et al.* 1994). Relatively large particles are likely to be retained in the pharynx and nasal cavity, while desorbed PAHs can be distributed to the tissues where biotransformation by phase I metabolic enzymes (i.e., cytochrome P450s) can occur. Collectively, the phase I and II metabolic pathways are responsible for converting PAHs to more water soluble metabolites and conjugating the metabolites to endogenous compounds (e.g., glutathione), which can then be excreted from the body (Castano-Vinyals et al. 2004, Shimada 2006). Although this mechanism prevents bioaccumulation, and thus has a protective role, some of the intermediates are known to be unstable and very reactive (Klassen 1996)

The initial step in PAH biotransformation is oxidation to oxides and phenols via cytochrome P450 isozymes; however, this activation pathway is cause for concern since the aforementioned diol-epoxides are formed by cytotochrome P450 (CYP)-catalysed reactions. CYP1A1 and 1B1, the primary enzymes responsible for activating carcinogenic PAHs in mammals, are present in many tissues, and expression of these enzymes is modulated by the binding of a PAH to the aryl hydrocarbon receptor (AhR). Figure 1.2 illustrates the mechanism by which CYP1A1 changes BaP to its tumourigenic bay-region diol epoxide.

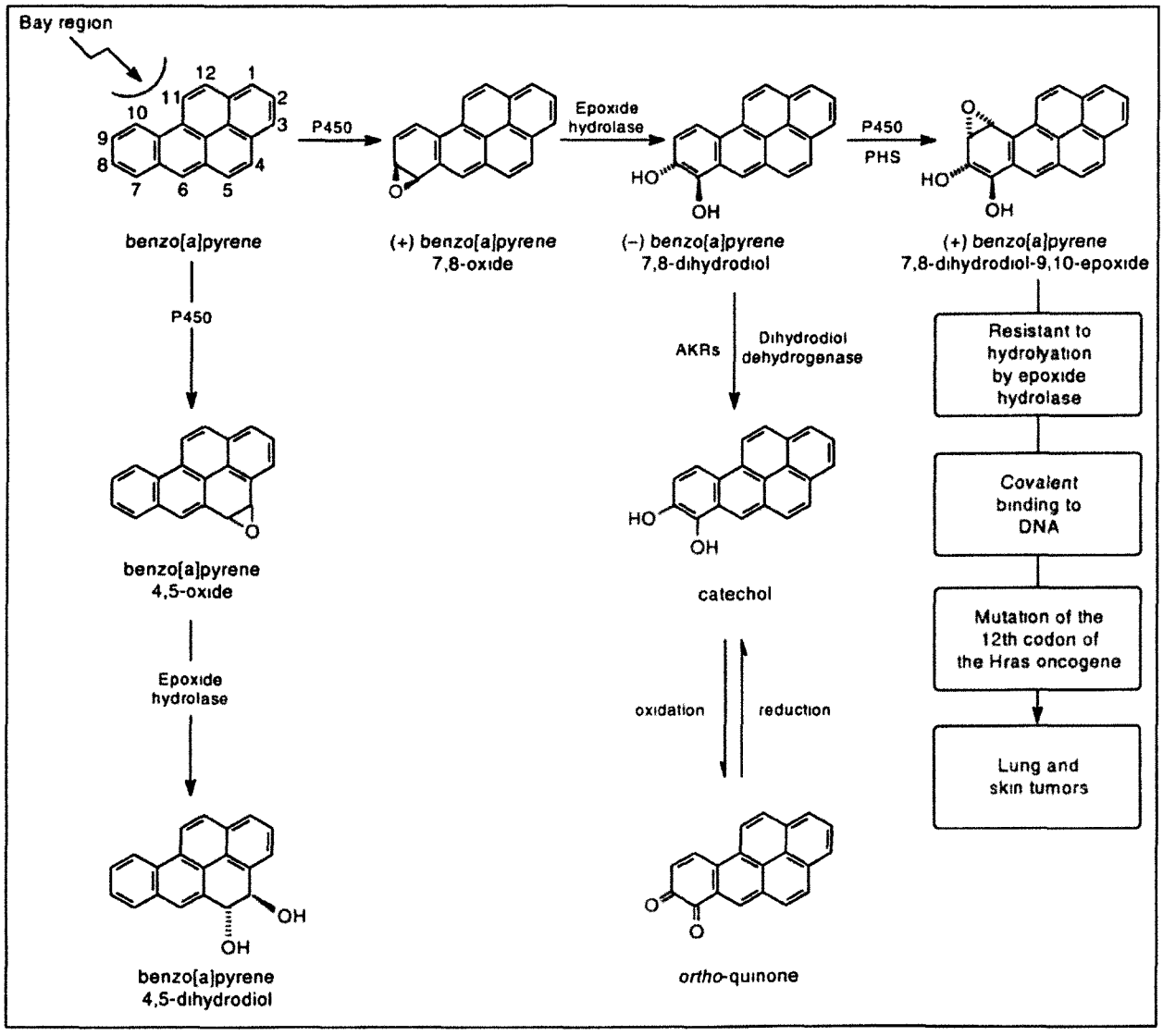


Figure 1.2 Metabolism of benzo[a]pyrene. Reactions catalyzed by Cyp1A isozymes include hydroxylation and epoxidation. Figure from Klassen (1996).

1.4 Coal Tar

Coal Tar is a brown-black liquid that is a by-product of the anaerobic pyrolysis of coal used to produce coke and/or coal gas. In the process of coking, the coal, which is heated to extremely high temperatures without oxygen, yields metallurgic coke that is used

as a blast furnace fuel during the production of steel (Scott et al., 2002). Before the 20th century coal tar had no commercial use and was often dumped in open pits, contaminating many coke oven and former manufactured gas plant sites. Once the commercial value of coal tar was realized, it was often sold to tar refiners, where it could be distilled into many different products including ingredients for wood preservatives, paints, pipe coatings, road-making material, and water proofing material. The exact composition of coal tar differs from batch to batch, but it is roughly composed of 48% hydrocarbons, 10% water, and 42% carbon.

Coal tar has been used since ancient times in the treatment of various skin related ailments. Two thousand years ago, Diascorides used a substance, which was then described as "asphalt" and closely resembles present day coal tar, to treat various cutaneous disorders. It was first demonstrated by Goeckerman in 1925 to be effective in the treatment of psoriasis in combination with ultraviolet radiation (Petrozzi *et al.* 1978, Pittelkow *et al.* 1981, Belsito & Kechijian 1982). This became known as the "Goeckerman regimen". Coal tar has a photosensitizing effect in the range of 330-550nm, and Goeckerman believed that it was the combination of coal tar and UVB that was so effective in treating psoriasis (Pathak & Joshi 1983). It was later found that UVB is not necessary in the treatment of psoriasis with coal tar (Thami & Sarkar 2002)

Coal tar remains an important substance to study, since it is common at orphaned contaminated sites, and a common ingredient in many consumer products. Products containing coal tar are still sold in pharmacies as topical treatments for psoriasis and head lice (DiSepio *et al.* 1999), and it is an ingredient in several pigments commonly used in hair dyes and makeup. These products exist in formulations containing coal tar concentrations

from 2% to 10%. Coal tar residues may also be present in the chimneys of homes heated with coal (ATSDR 2002).

Exposure to coal tar via coke oven emissions is a concern for workers in the aluminum, graphite, steel, electrical and construction industries. Coal tar, coal tar pitch, and coal tar pitch volatiles are used in roofing and road paving products. Studies of coke oven workers who have been exposed to coke oven emissions, as well as tar distillery workers and roofers, have reported an increase in cancer of the lung, trachea, bronchus, kidney, and prostate (Swaen *et al.* 1991, Swaen & Slangen 1997). Chronic exposure to these emissions has also been shown to cause conjunctivitis, severe dermatitis, and lesions of the respiratory and digestive system (U.S EPA, 1999).

Coal tar and coal tar extract yields a positive response in a vast array of assays for mutagenicity and clastogenicity, and is also known to be carcinogenic in experimental animals (IARC 1983, Mukhtar *et al.* 1986, Chang *et al.* 1992). It is also known to initiate skin tumors when applied topically to the skin of mice (Mukhtar *et al.* 1986). However, studies that have looked at the effects manifested in humans following dermal treatment for skin ailments are conflicting. A 25 year study by Muller and Perry (1984) noted that there was no significant increase in cancer incidence in patients treated with coal tar (Muller & Perry 1984). However, Grupper et al. noted that there was 0.2% increase in skin cancer incidence in 2393 psoriasis patients treated with topical coal tar (Grupper 1971). The conflicting results of these studies might be explained by the differences in the composition of the coal tar mixtures, which are not standardized. The absence of cancer findings in the Muller and Perry study may also be explained by interactive effects of the PAHs in the complex mixture. For example, it has been shown that PAHs present complex mixtures can show decreased mutagenic activity (Lemieux *et al.* 2008). Other studies have

documented antagonistic interactions between specific PAHs. For example, mixtures of BaP and 5-methylchrysene can inhibit the formation of their own metabolites by competitive inhibition of CYP1A1 and 1B1 (Shimada 2006, Donnelly *et al.* 1998). Such phenomena will decrease the amount of carcinogenic metabolites formed (Mahadevan *et al.* 2007), thus accounting for the decreased activity of PAHs in complex mixtures such as coal tar, relative to what would be expected based on the sum of the expected contributions from the individual PAHs.

The two major routes of coal tar exposure are dermal and inhalation. If the environmental concentration of coal tar is high, oral exposure also becomes relevant, especially for children playing at or near contaminated areas. In spite of its ongoing and frequent use in dermatology, the pharmacologic modes of action of coal tar extract are still not known. It has been suggested that it has antibacterial, antiparasitic, antipruritic and antifungal properties (Muller & Perry 1984), however, there is a paucity of data to support these claims.

1.5 Muta™Mouse FE1 Pulmonary Epithelial Cells

In vitro assay systems are considerably less expensive and laborious than *in vivo* systems. Consequently, this project, which constitutes a portion of the "proof of principle" for the use of a combined genomic and proteomic approach to discover useful biomarkers for complex mixtures such as coal tar, employed an *in vitro* system based on transgenic murine cells.

The Muta™Mouse FE1 lung cell line was selected since it is a well-characterized transgenic cell line that has been employed to analyse a variety of mutagenic substances, including BaP and coal tar, and moreover, has been found to be a suitable complement to the *in vivo* Muta™Mouse mutation assay (White *et al.* 2003, Berndt-Weis *et al.* 2009). In addition, since inhalation is one of the routes for coal tar exposure, and lung can be a target organ, FE1 is an appropriate choice for this project. A major advantage of this particular cell line is its λgt10*lacZ* shuttle vector, which contains 40 concatenated copies of the *lacZ* gene for mutation scoring (Gossen *et al.* 1989). Following an exposure of interest, the transgene can be recovered from genomic DNA using a λ bacteriophage packaging system, and scored using a bacterial positive selection system. The transgene assay used in this thesis is more thoroughly outlined in section 2.3.5.

1.6 Objectives and Hypothesis

This project constitutes the initial stages of the work required to identify a unique biomarker (e.g., gene expression or proteomic profile) of coal tar exposure. Gene expression profiling will be carried out on murine epithelial cells exposed to increasing concentrations of coal tar selected based on cytoxicity and mutagenicity. The resulting expression profiles will be employed to discern mechanisms of action of coal tar extract, and these mechanisms will be discussed. In addition, identified genes that code for secreted protein will be discussed in relation to their utility as biomarker candidates.

Hypothesis:

The exposure of murine pulmonary epithelial cells (i.e, FE1 cells) to coal tar extract will elicit changes in gene expression. The the observed changes will provide mechanistic

insight into the toxicological properties of coal tar extract, and moreover, reveal candidate biomarkers of exposure.

Chapter 2: Materials and Methods

2 Materials and Methods

2.1 Cell culture

Muta™Mouse FE1 pulmonary epithelial cells were propagated from one vial of cryo-preserved passage 11 cells. Cells were grown on 150 mm plates (Gibco-Invitrogen, Burlington, Ontario), in a 1:1 mixture of Dulbecco's Modified Eagle's medium and F12 nutrient mixture (DMEM /F12) supplemented with 100 U/ml penicillin G, 100 mg/ml streptomycin, (Gibco-Invitrogen, Burlington, Ontario), 2% (v/v) fetal bovine serum (FBS), and 1 ng/mL murine epidermal growth factor (EGF) (Gibco-Invitrogen, Burlington, Canada). All incubations were carried out at 37 °C, 95% humidity, and 5% CO_2. The cells were sub-cultured every 48 hours in order to provide sixteen 150 mm plates, per replicate, of 70% confluent cultures within 10 days. The percent confluence was determined using replicate counts on a Coulter Particle Counter (Beckman Coulter, USA). Twenty four hours before the exposure (day 9), cells were washed with phosphate buffered saline (PBS) and serum-free medium (SFM) was applied. On the 10^{th} day, passage 14, 160 dishes of FE1 cells were ready for exposure, 60 of which were used for this project. See Figure 2.1 for schematic showing the process used to generate sufficient FE1 cells.

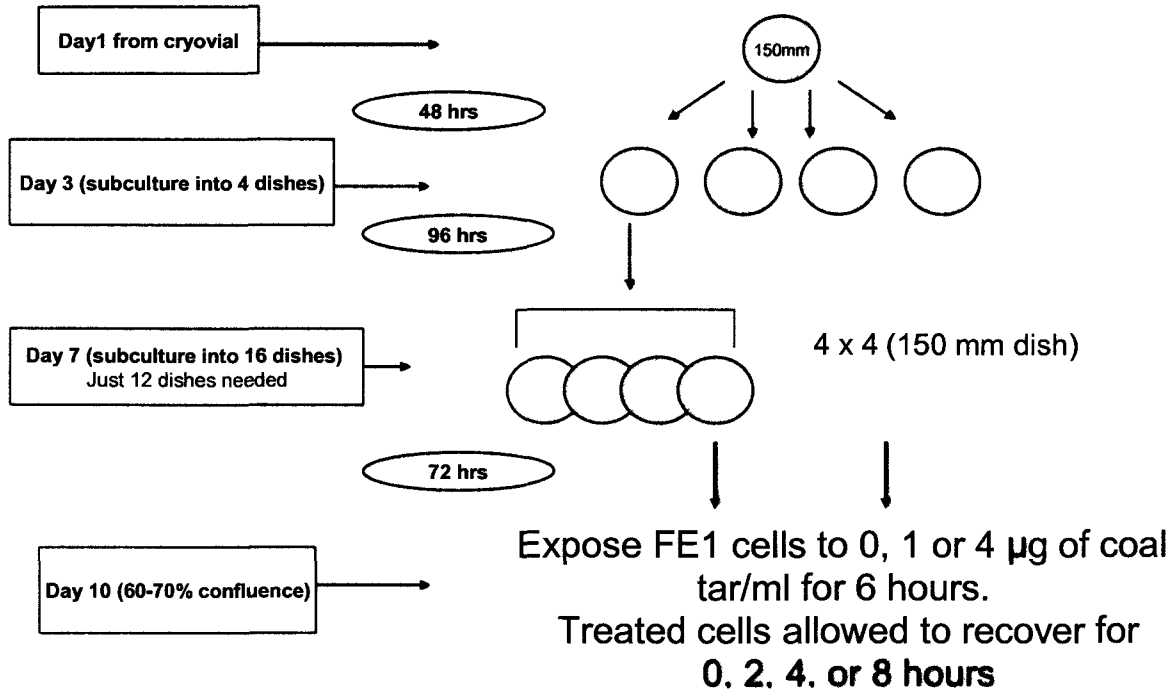


Figure 2.1 Schematic showing the processing of FE1 cells from passage 11 to passage 14. This process was repeated 5 times for the 5 replicates.

2.1.1 Standard Reference Material 1597a

The experiments conducted in this thesis employed Standard Reference Material (SRM) 1597a (i.e., coal tar extract) as a standardized complex mixture of PAHs. SRM 1597a, a combustion-related mixture of polycyclic aromatic hydrocarbons isolated from coal tar, is supplied as a toluene solution, and it can readily be obtained from the National Institute of Standards and Technology (National Institute of Standards and Technology). The material used to prepare SRM 1597 was obtained from a medium-crude coke oven tar. The coal tar sample was processed through a liquid chromatographic column containing attapulgus clay to remove the highly polar components. 10% methylene chloride in n-pentane was used to elute the sample, which was concentrated until a thick orange syrup remained. This material was then dissolved in 4.5 L of toluene and aliquoted into 5 mL amber ampoules. Each ampoule contains approximately 1.3 mL of solution (Certificate of Analysis, Standard Reference Material 1597a). The concentration of PAHs in SRM 1597a is listed in Table 1.5.

Table 2.1 Certified Concentrations of 16 priority PAHs in SRM 1597a
Error values represent the standard deviation

PAH	Mass Fraction (mg/kg)
Naphthalene	1030 ± 100
Acenaphthylene	263 ± 7.0
Acenaphthene	7.63 ± 0.26
Fluorene	145 ± 4.0
Phenanthrene	454 ± 7.0
Anthracene	107 ± 3.0
Fluoranthene	327 ± 7.0
Pyrene	240 ± 7
Benzo[*a*]anthracene	98.1 ± 2.3
Chrysene	66.2 ± 5.3
Benzo[*b*]fluoranthene	66.1 ± 4.4
Benzo[*k*]fluoranthene	41.2 ± 0.4
Benzo[*a*]pyrene	93.5 ± 1.4
Benzo[*g,h,i*]perylene	50.5 ± 0.6
Indeno[1,2,3-*c,d*]pyrene	55.5 ± 0.8
Dibenzo [*a,h*]anthracene	6.93 ± 0.40

2.2 Coal Tar Extract Preparation

Coal tar extract (i.e., SRM 1597a) was originally dissolved at 8 mg coal tar/ ml of toluene. The toluene was removed under a stream of ultra-pure nitrogen, and replaced with DMSO. Coal tar extract dissolved in DMSO was stored at room temperature in the dark until used for the exposure outlined in section 2.4.

The results of cytoxicity and mutagenicity analyses were used to choose the coal tar exposure concentrations (see results section 3.1 for details). The concentrations of coal tar employed were 1 and 4 μg of coal tar /ml with a final DMSO concentration of 1% (v/v). According to the data presented in Table 1.1, this concentration is equivalent to 3.052 ng priority PAHs per ml. The control received 1% DMSO in SFM.

2.3 Cytotoxicity and Mutagenicity Assays to Choose Exposure Concentrations

This preliminary phase of the project was designed to identify concentrations that would likely elicit gene expression changes, but not kill the cells. The endpoints selected to assess cytotoxicity, cell viability and mutagenicity are displayed in Table 2.1.

Table 2.2 Mutagenicity and cell viability/cytotoxicity endpoints employed to select coal tar extract concentrations.

Assay	**Endpoint**
Cell enumeration	• Cell viability, growth and replication.
Trypan Blue exclusion	• Cell viability
LDH release	• Membrane integrity and cell viability
XTT reduction	• Mitochondrial activity
Muta™ Mouse transgenic mutation assay	• Gene mutation

2.3.1 Cell Enumeration

Cells (70% confluent) were exposed to several concentrations of coal tar extract for 3 hours, with no recovery time. Treated cells were then trypsinized, and a sub-aliquot used to assess cell numbers. The following concentrations were investigated 0, 8 x 10^{-5}, 16 x 10^{-5}, 20 x 10^{-5}, 80 x 10^{-5} and 800000 x 10^{-5}µg coal tar/ml.

2.3.2 Cell Viability Assessment via Trypan Blue Exclusion

The Trypan blue exclusion cell viability assay assesses the percentage of viable cells by examining the proportion of the cell population that has an intact membrane, and is capable of blocking or excluding the entrance of the dye. Dead or dying cells readily take up the Trypan Blue, which is apparent under magnification (Jauregui *et al.* 1981).

Cells from passage 14 were grown for 72 hours, rinsed with PBS, incubated in SFM for 24 hours, and exposed to coal tar for 6 hours. Solvent controls were exposed to DMSO for 6 hours. A hemocytometer was loaded with 20 µl of cells in suspension, dyed with an equal volume of trypan blue (0.4% (w/v) in PBS) for 2 minutes, and the number of stained cells counted using phase contrast microscopy. The percentage of viable cells relative to the control was calculated using the following equation:

Percent Viable cells = Total viable cells (unstained)/ total cells (stained and unstained) x 100

The coal tar concentrations used for the Trypan Blue Exclusion assay were 0.8 x 10^{-5}, 16 x 10^{-5}, 20 x 10^{-5}, 80 x 10^{-5} and 800000 x 10^{-5}.µg coal tar/ml.

2.3.3 Lactate Dehydrogenase (LDH) Release Cytotoxicity Assay

The LDH Release assay is a cytoxicity assay that employs the ability of LDH (lactate dehydrogenase) to catalyze the reduction of pyruvate, yielding lactate and oxidized NAD+ (Aly & Domenech 2009, Bonfoco *et al.* 1995). The oxidization of NADH is spectrophotometrically measured by a decrease in the absorbance at 490 nm (i.e., A_{340}) (Thermo Scientific, Waltham, MA). Like the Trypan Blue exclusion assay, it is also assess membrane integrity, since disruption of membrane integrity causes release of LDH. The LDH release assay has been shown to be a good indicator of cellular damage and toxic stress in many different cell lines (Legrand *et al.* 1992, Aly & Domenech 2009).

An LDH assay kit was obtained from Aniara Corporation (Mason, Ohio, U.S). Passage 17 cells were used to seed 96 well plates with approximately 40,000 cells/well, and plates were incubated for approximately 48 hours. Cells were rinsed with PBS, SFM added to each well and the cells incubated for 24 hours. The plate was then centrifuged, SFM discarded, and the cells exposed to the desired coal tar extract concentration (in SFM). Control wells were treated with SFM containing DMSO. Cells were exposed for 3 or 6 hours. Eight concentrations were chosen for the 6 and 24 hour exposure 0, 0.2, 0.5, 1.0, 2.0, 4.0, 6.0 and 8.0 μg/ml.

2.3.4 XTT Formazan Cytotoxicity Assay

The XTT assay assesses mitochondrial function by measuring dehydrogenase activity. The crucial component of the XTT assay is the sodium salt of (2, 3-bis (2-methoxy-4-niro-5-sulfophenyl)-2H-tetrazolium-5carboxyanilide. Mitochondrial dehydrogenase activity is assessed spectrophotometrically via the appearance of soluble orange formazan crystals. The degree of cytotoxicity is positively correlated with the amount of colour detected spectrophotometrically at 500nm (Roehm *et al.* 1991).

An XTT assay kit was obtained from Xenometrix (Allschwil, Switzerland). Passage 17 cells were used to seed 96 well plates with approximately 40,000 cells/well. Cells were incubated for approximately 48 hours, rinsed with PBS, and incubated for a further 24 hours in SFM. The plate was then centrifuged, the SFM discarded, coal tar extract (in SFM) added, and the cells exposed for 6 hours or 24 hours. Serum free media with DMSO was added to the control wells. The XTT reagent was added to the exposed cells according to the manufacturer's instructions. The absorbance at 450 nm was measured using a spectrophotometer (Thermo Scientific, Waltham, MA). Coal tar concentrations for the 6 and 24 hours exposures were 0, 0.2, 0.5, 1.0, 2.0, 4.0, 6.0 and 8.0 μg/ml.

2.3.5 Muta™Mouse Mutagenicity Assay

The FE1 cell line selected for this project contains an integrated lacZ transgene that can be used to score mutations. Scoring of lacZ transgene mutations employed the phenyl-ß-D-galactopyranoside (P-gal) positive selection system described in Vijg and Douglas (1996), and the method employed to score transgene mutations in FE1 cells is described in Lemieux et al., 2006 (Lemieux 2006). Briefly, exposed mouse cells are lysed, genomic DNA is packaged into lambda phage heads, galE- host bacteria (E coli) are infected with the rescued DNA containing the transgene, and mutant transgenes are scored as plaques formed in the presence of P-gal, the selective agent. Total plaque-forming units (pfus) are scored in the absence of the selective agent, and the lacZ mutant frequency is calculated as the ratio of the mutant plaques to total pfus. A general overview of this assay is illustrated in Figure 2.2.

FE1 cells (passage 14) were exposed to coal tar extract for 6 hours, and incubated for 72 hours for mutation fixation. The concentrations of coal tar extract used for this assay were 0, 1.05, 2.1, 8.38, 10.48, 12.58, 20.96 μg of coal tar/ml.

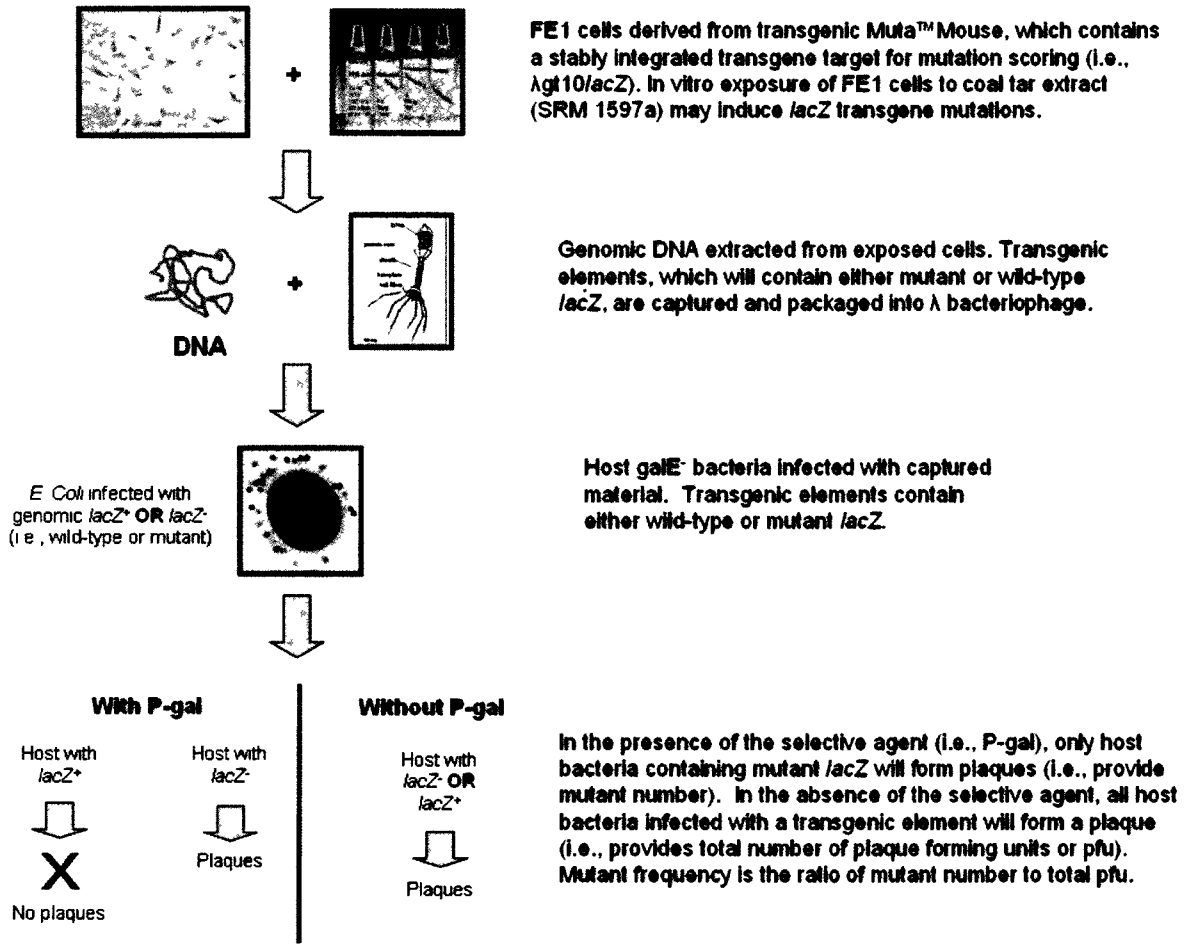


Figure 2.2 General overview of the FE1 Muta^{TM}Mouse Mutagenicity Assay

2.3.5.1 DNA extraction, handling and *lacZ* transgene scoring

FE1 cells were rinsed with PBS and lysed overnight (10 mM Tris pH 7.6 (Caledon Laboratories Ltd, Georgetown, ON), 10 mM EDTA (Sigma-Aldrich Canada Ltd, Oakville, ON), 150 mM NaCl, 1% (w/v) SDS and 1mg/ml proteinase K). The lysate was collected, and the DNA was purified using phenol:chloroform extraction followed by phenol:chloroform:isoamyl alcohol extraction. The DNA was precipitated in an ethanol solution containing 200 mM NaCl, and spooled onto a sealed Pasteur pipette. The collected DNA was rinsed in ethanol, air-dried, and dissolved in 100 µl of TE buffer (10 mM Tris pH 7.6, 1 mM EDTA). The DNA was stored at 4°C until further use.

Transpack™ was used to package the extracted genomic DNA into λ phage particles (Stratagene, La Jolla, CA). The phage particles containing the packaged DNA were mixed with the host bacterium (*Escherichia coli* Δ*lacZ*, *gal*E$^-$, *rec*A$^-$, pAA119 with *gal*T and *gal*K), plated on minimal agar with 0.3% w/v P-gal, and incubated overnight at 37°C. A titre plate without P-gal was completed in tandem. Plaques were manually scored the following day and mutant frequency calculated as the ratio of mutant plaques and total plaque forming units (pfus).

Poisson regression was employed to analyze the mutant frequency results using SAS 9.2. The data were fit to the model $\log(E(Y_i)) = \log t_i + ß\mathbf{x}_i$, where $E(Y_i)$ represents the expected value for the ith observation, ß represents the vector of regression coefficients, t_i is the offset variable to explain the differences in observation count. This was assigned a constant coefficient of 1.0 for each observation. $\mathbf{x}_i$ is a vector for the covariate of the ith

observation. The Holm-Bonferroni correction for multiple comparisons was employed for post hoc contrasts between exposed and control.

2.4 Coal Tar Extract Exposures for Gene Expression Analyses

Five biological replicates were exposed to one of two concentrations: 1 and 4 μg of coal tar per ml (3.052 ng of PAH/ug of coal tar), in 25 ml of SFM containing a final concentration of 1% DMSO (v/v). The treated cells were incubated for 6 hours, washed twice with PBS, and allowed to recover in SFM for 0, 2, 4 or 8 hours. Cells were then treated with TriZol® LS reagent (Invitrogen Life Technologies, Carlsbad, CA) and stored at -80°C until RNA extraction. Five replicates were used for each condition, totalling 60 samples (3 concentrations x 4 recovery time points x 5 replicates).

2.5 RNA Extraction

Total RNA from five replicates was harvested using TriZol® LS reagent (Invitrogen Life Technologies, Carlsbad, CA) and RNEasy Mini columns (Qiagen, Mississauga, ON). RNA was quantified using a UV Spectrophotometer. All RNA samples showed A_{260}:A_{280} ratios between 1.9 and 2.1, and RNA quality was verified using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Only high-quality RNA samples (i.e., 28S/18S > 1.8) were used for further analysis.

2.6 Microarray Analyses

Agilent high density murine oligonucleotide 22K microarrays were used for this study (Agilent Technologies Inc., Mississauga, ON). Control and coal tar exposed RNA samples were fluorescently labeled with cyanine 5 (Perkin-Elmer Life Sciences,

Woodbridge, ON), and universal reference mouse total RNA (Stratagene, Cedar Creek, TX) was labeled with cyanine 3 using an Agilent Low Input Linear Amplification Kit, according to the manufacturer's instructions. Briefly, double-stranded cDNA was synthesized using retroviral MMLV reverse transcriptase with T7 promoter primer, starting with 2 μg total RNA. The labeled cDNAs were then transcribed into cRNA using T7 RNA polymerase. The amplified cRNA was precipitated with anhydrous ethanol and purified with RNEasy mini columns (Qiagen, Mississauga, ON) following the manufacturer's instructions. The cleaned cRNA was eluted with RNase-free water following centrifugation for 30 seconds at 13,000 rpm. The quantity and quality were measured using a UV spectrophotometer. Five micrograms each of labeled Cy3 and Cy5 amplified cRNA were used to make the probe. Control targets, target solution and 25x fragmentation buffer were mixed with the probe in a 1.5ml nuclease-free micro-centrifuge tube and incubated at 60°C for 30 minutes, and hybridization buffer was added to the probe. The samples were manually pipetted onto a gasket slide and the 60-mer oligonucleotide microarray was placed, face down, on top of the gasket slide. The samples were then incubated overnight (17 hours) at 60°C. Microarrays were scanned using a Scanarray Express (Perkin-Elmer Life Sciences, Woodbridge, ON) at various laser intensities depending on the intensity of the fluorescing probes. Digitized intensity differences in the hybridization patterns were interpreted as copy number differences between the test and reference genomes using image analysis software (Krane 2004).

Imagene 5.5 (BioDiscovery, Inc., El Segundo, CA) was used to quantify microarray image spots. Gene spots were identified as present if their signals were greater than the mean of the negative control plus three times the standard deviation. The data were

analyzed using a reference design with median signal intensities. The background fluorescence, which was assessed using a negative control spot, was used to identify poor quality arrays (i.e., arrays with high background). Spots with median signal intensities less than three trimmed standard deviations below the trimmed mean were also excluded. The data were normalized using the Lowess method, and ratio intensity plots were inspected for quality control. Hierarchical clustering dendrograms of reference versus control samples indicated that there were no outliers within the 58 samples. Box plots (see Appendix B) and signal intensity plots were constructed and inspected for all normalized and pre-normalized data. A microarray analysis of variance (MAANOVA) was used to identify significant differences between control and exposed cells for individual concentrations within times and their interactions, and p-values for all the statistical tests were determined using the permutation method. These p-values were adjusted for multiple comparisons by applying the False Discovery Rate (FDR) statistic (FDR adjusted p-values are reported in Table 3.1 and 3.2). Two microarrays were excluded due to low signal intensities and two others were excluded based on clustering. See Appendix A for a more detailed description of the statistical methods used.

2.6.1 Condition Trees

Normalized data were imported into Genespring 7.0 (Agilent Technologies, Mississauga, ON), and condition trees were created to illustrate patterns of expression according to the magnitude and direction of fold changes. Condition trees were created to illustrate the expression across all conditions, all significant genes (FDR adjusted $p<0.05$), significant genes from early recovery time points (0 and 2 hours recovery, FDR adjusted

$p<0.05$), and significant genes from the late recovery time points (4 and 8 hours recovery, FDR adjusted $p<0.05$).

2.6.2 Gene Ontology (GO) Analysis

A Gene Ontology analysis was completed using DAVID Bioinformatics Resources (National Institute of Allergy and Infectious Diseases, NIH, Frederick, MD) (Huang da et al. 2009, Dennis et al. 2003) and Genespring 7.0 (Agilent Technologies). Several gene lists were analyzed using these tools. Two gene lists were compiled based on criteria of varying stringency, and two separate GO analyses conducted. In order to ensure that all important gene ontology motifs were captured, one ontology analysis included genes that met one of three criteria :(1) significantly changed (FDR adjusted $p<0.05$), (2) significantly changed with unadjusted $p<0.001$, or (3) greater than 2-fold change in expression. The second ontology analysis only included genes that showed significant change after the false discovery rate adjustment was applied (FDR adjusted $p<0.05$)

Once the ontological analysis was completed, the large list of gene ontology categories was narrowed down by including only categories that were deemed significant by DAVID, after adjustment for false discovery (i.e.FDR adjusted $p<0.05$). Only ontology groups with 5 or more genes were included in the final results, and any redundancy was removed. Ontology categories were also removed if they were considered too general to be informative (e.g., "Cellular process"), or were not considered biologically relevant for FE1 lung cells (e.g., motifs relating to brain function or eyesight). The same analyses were repeated using Genespring 7.0 (Agilent), and the results were combined.

2.6.3 Pathway Analysis

The two groups of genes used in the ontology analyses (see section 2.6.2) were reanalyzed using an object search in the Kyoto Encyclopedia of Genes and Genomes (KEGG). The pathway analysis using the combined gene lists was reduced in size by excluding any pathway containing less than 5 genes showing altered expression. The pathway analysis that only included the genes showing FDR adjusted $p<0.05$ was less restrictive, and pathways containing 2 or more genes were retained.

Chapter 3: Results

3 Results

3.1 Cytotoxicity and Mutagenicity

Cytotoxicity and mutagenicity assays were used to choose effective coal tar exposure concentration for subsequent gene expression analyses (i.e., 1 and 4 μg coal tar/ml). Since coal tar extract is known to contain a variety of potent mutagens, the *lacZ* transgenic mutagenicity assay heavily influenced the concentration choices. The other cytotoxicity assays were used to ensure that the concentrations examined would not induce extensive mortality. It was decided that the primary concentration (i.e., low concentration of 1 μg/ml) should have at least 60% cell viability, or show relatively little cytotoxicity. The cell viability, as determined by cell enumeration was low for both selected concentrations; however, the cytotoxicity assays indicated that the cells are likely sufficiently impacted to elicit interesting gene expression responses. The high concentration resulted in higher, but nevertheless acceptable level of cell mortality for both cell viability assays.

3.1.1 Viability Assessment by Cell Enumeration

The cell enumeration and Trypan Blue exclusion assays were employed as range-finding assays to determine the active concentration range for coal tar extract (i.e., ng per ml versus μg per ml or mg per ml). The results of the post-exposure cell enumeration analyses are illustrated in Figure 3.1. The results from this assay were in close agreement with the results obtained using the Trypan Blue Exclusion assay.

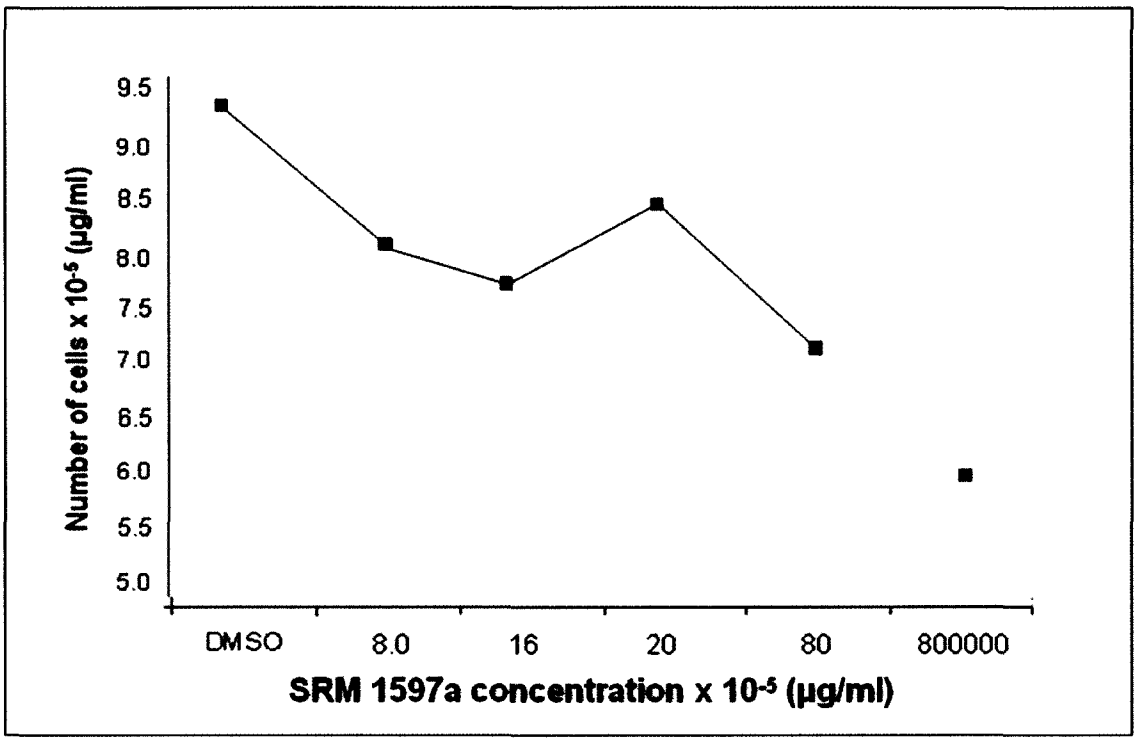


Figure 3.1 Viability of FE cells exposed to coal tar extract, by post-exposure cell enumeration.

3.1.2 Trypan Blue Assay

Figure 3.2 illustrates the results of the Trypan Blue Exclusion Assay. The results indicate that cell viability remained unaffected by the coal tar concentraton up to 0.8 ng/ml.

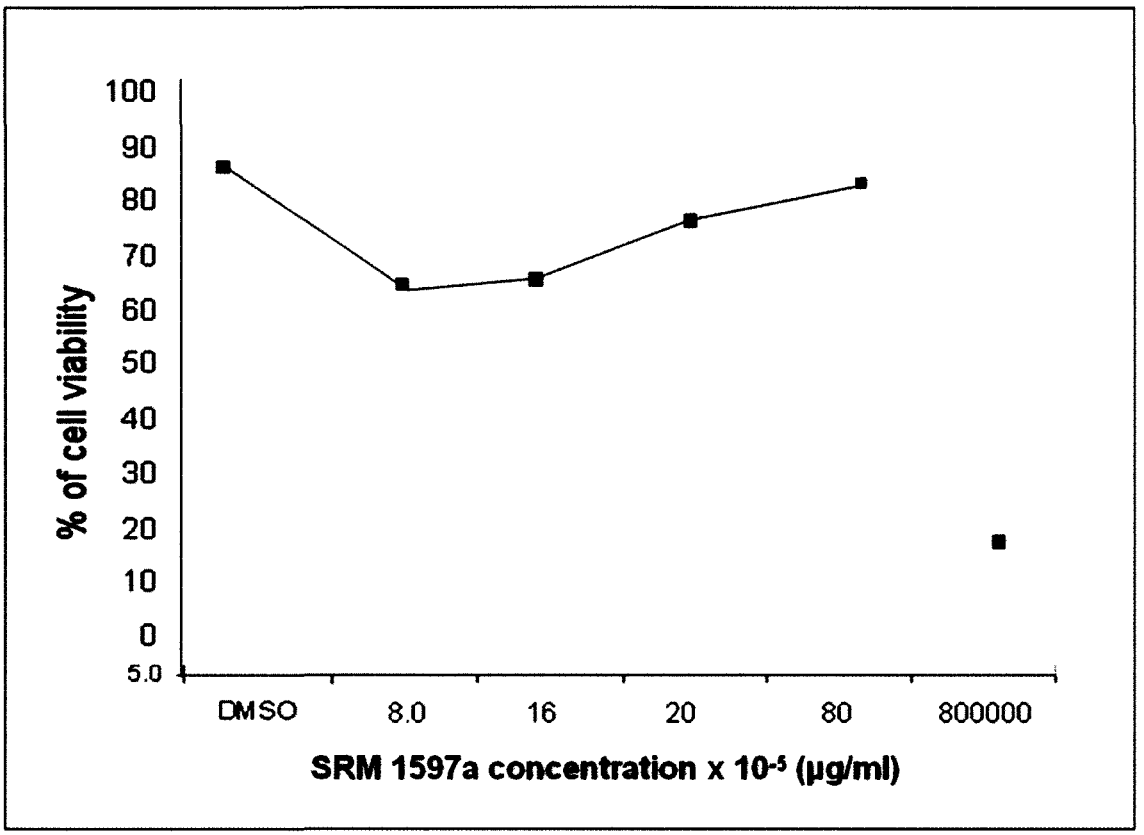


Figure 3.2 Cytotoxicity of coal tar extract assessed using the Trypan Blue exclusion method. Cells were counted using a haemocytometer examined by phase contrast microscopy.

3.1.3 Lactate Dehydrogenase (LDH) Release Assay

The results of the LDH release assay are illustrated in Figure 3.3. As expected, the 6 hour exposure showed less cytotoxicity in comparison with the 24 hour exposure. Both of the selected concentrations, i.e., 1 and 4 μg/ml, showed relatively little difference in membrane integrity compared to control, as determined by measurement of LDH leakage.

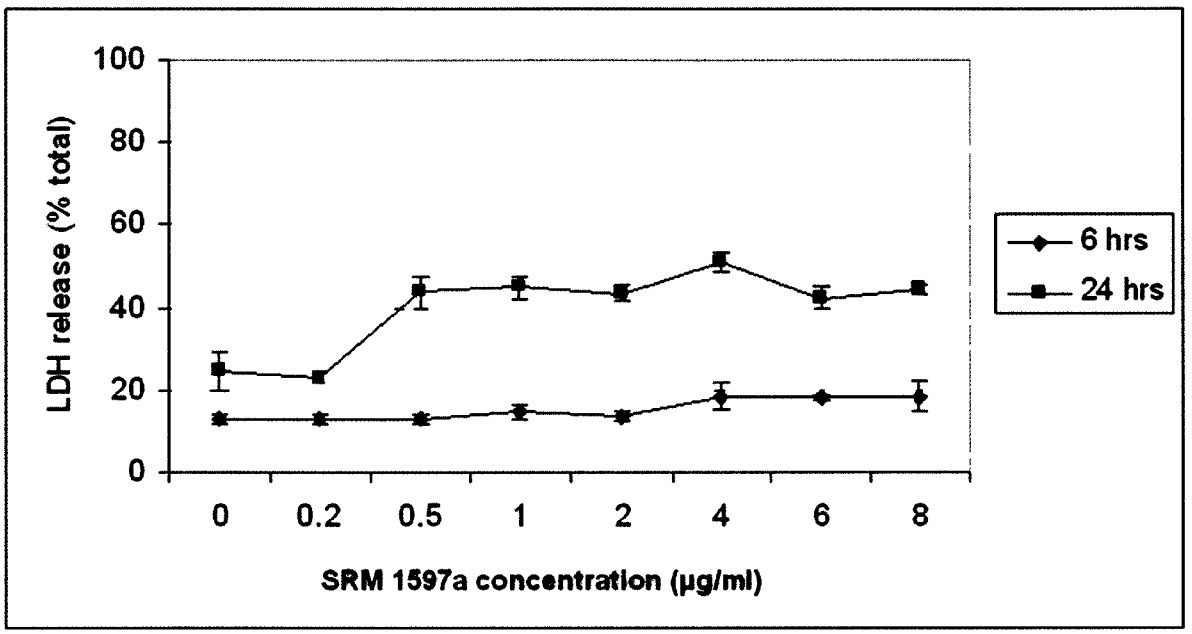


Figure 3.3: Cytotoxicity of coal tar extract assessed using the LDH leakage assay. FE1 cells were treated with increasing concentrations of coal tar extract for 6 and 24 hours, and mean cell viability values are expressed as a percentage of control. Error bar represent one standard error of the mean.

3.1.4 XTT Reduction Assay

The results of the XTT assay are illustrated in Figure 3.4. Both the 6 and 24 hour exposure show a downward trend over time, however, the 24 hour data, unexpectedly, indicates that the mitochondrial dehydrogenase activity was less affected relative to that observed at 6 hours.

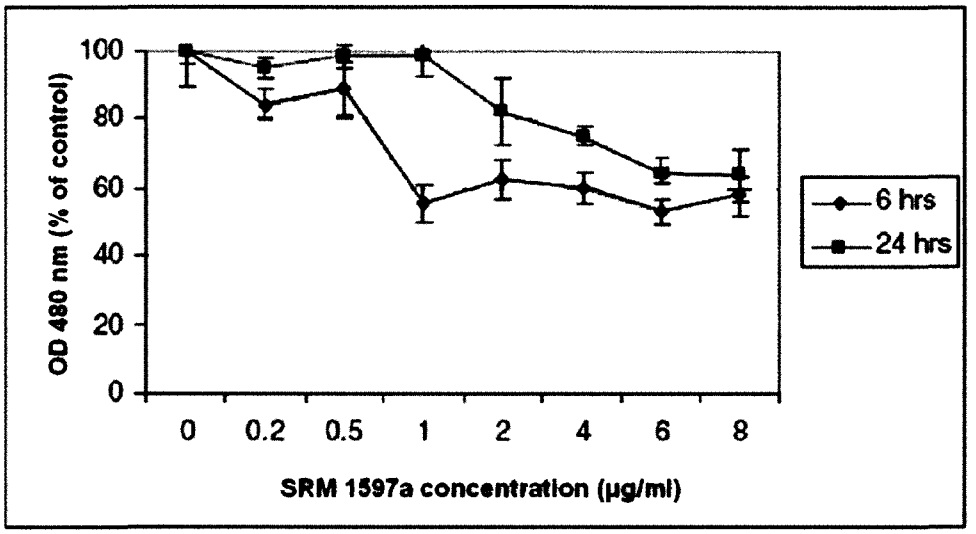
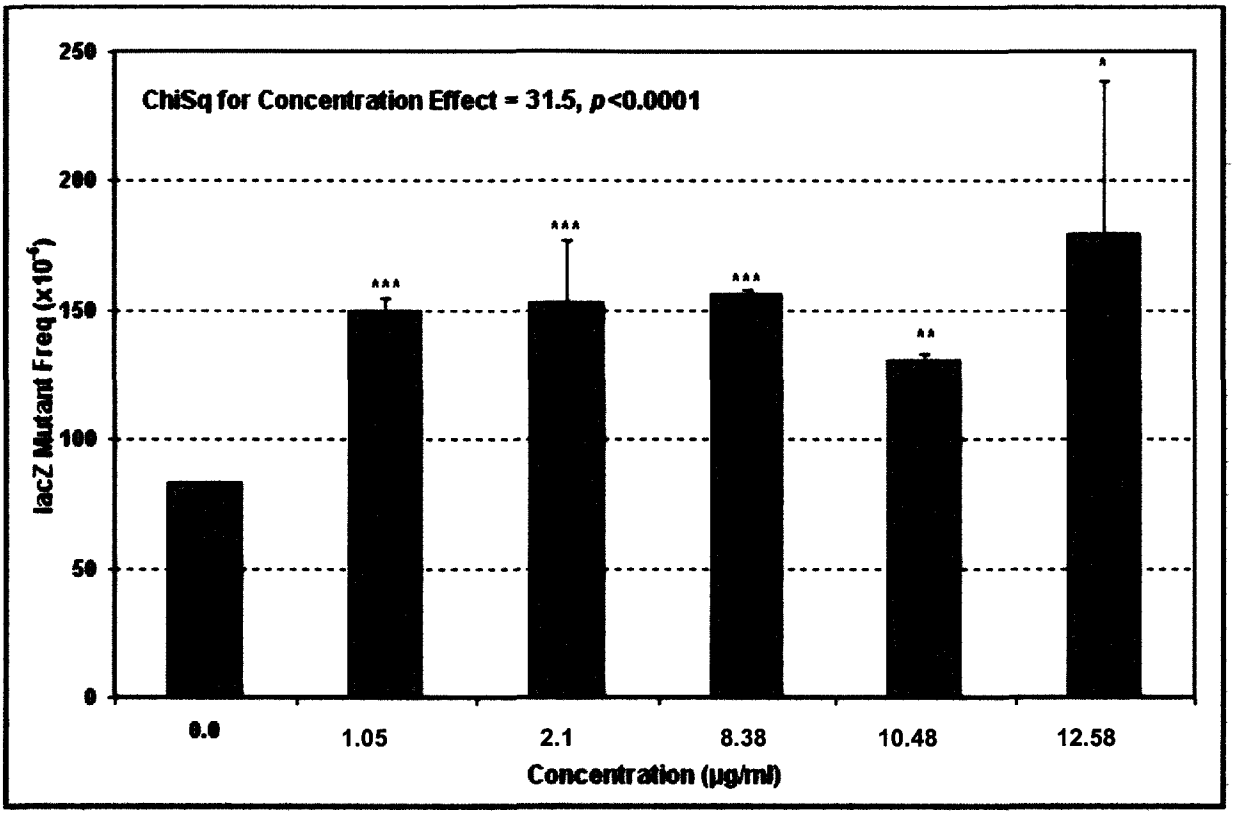


Figure 3.4: Cytotoxicity of coal tar extract assessed using the XTT reduction assay. FE1 cells were treated with increasing concentrations of coal tar extract for 6 and 24 hours and mean cell viability values are expressed as a percentage of control. Error bars represents one standard error of the mean.

3.1.5 Muta™Mouse FE1 Mutagenicity Assay

The results of the *lacZ* transgene mutagenicity assay are illustrated in Figure 3.5. These results heavily influenced the selection of concentrations for the subsequent gene expression analyses. Poisson regression analysis indicates that most of the tested concentrations induced a significant increases in *lacZ* mutant frequency compared to control (i.e., $p<0.05$). The concentrations chosen for the microarray experiment, 1 and 4 µg coal tar/ml, represent concentrations with high mutagenicity. Higher concentrations that yielded a substantial increase in mutant frequency relative to control (e.g., 13 µg/ml), were

not selected for gene expression analyses since concentrations above 4μg/mL are associated with higher cytotoxicity.



*significant p<0.05

**significant p<0.01

***significant p<0.001

Figure 3.5 Mutagenicity of coal tar extract assessed using the *lacZ* transgene mutation assay in Muta™ Mouse FE1 cells. Note – 1µg of coal tar ~ 3.05 ng priority PAHs.

3.2 Microarray Analyses

3.2.1 Global gene expression

Glogal gene expression changes were assessed for cells exposed to two concentrations of coal tar extract at 4 recovery time points. The results from the global gene expression analyses were broken up into two groups, early recovery time points, including 0 and 2 hours, and late recovery time points, including 4 and 8 hours. This was done due to unexpectedly high expression of *Cyp 1b1* in the control samples, indicating that a sample had likely been switched within the early time points. Two microarrays were excluded due to low signal intensities and high background levels, and two others were excluded based on clustering results (This analysis identified two samples that had been mislabelled, which was apparent by the abberent expression of 2 samples which had been labeled labeled as controls, but had clustering patterns of the treated conditions). Genes were considered 'present' if the spots were greater tha 3 trimmed standandard deviations above the trimmed mean of the negative control spots on the array. The spots had to meet this criteria in 10 samples all together to be considered present. Out of the early time points 2,960 were absent and 18,359 were present. Amongst the late time points 5,295 were absent and 16,103 were present. Twenty six and 102 genes in the early and late recovery time point groups, respectively, were identified as significantly different from the control at FDR adjusted $p<0.05$ in at least one experimental condition. Seven genes from the late time points list were not known genes (as of September 2009), and could therefore not be annotated. Therefore, subsequent ontology and KEGG analysis did not include these genes. The complete lists of differentially expressed genes are provided Appendix C Table

5.1 (early time points) and Table 5.2 (late time points), respectively. There were 17 genes in common between these two lists. The main effects of the factorial ANOVA indicate that concentration had the greatest proportion of variation explained by the treatment (ie the highest amount of genes associated with the concentration main effect). In contrast, there were very few genes that were associated with a significant result from the other 2 main effects (recovery time and concentration-recovery time interaction). Genes that have the top 10 fold-rank change (i.e. The 10 most differentially expressed genes) for both early and late time points, are displayed in Table 3.2 and Table 3.3 respectively.

The largest fold change in the late time points was observed for *Cyp1a1,* which was upregulated 40- and 34 fold after 8 hours of recovery for 1 and 4 μg/ml of coal tar treatment, respectively. For the 4 hour recovery time point, *Cyp1a1* showed a 23- and 14 fold increase. The high fold changes for *Cyp1a1* were not seen in the early time points, and this is likely due to data quality issues (i.e., a sample was likely switched, dampening the treatment effect).

The majority of the genes identified as significant (i.e., FDR adjusted $p<0.05$), were up regulated compared to the control. See Table 3.1 for a breakdown of the number of genes found to be up regulated, versus down regulated for each pairwise comparison (i.e., treated versus control). Most of the significant pairwise comparisons were up-regulated with respect to control. There were 77 upregulated genes and 33 down-regulated genes that showed a statistically significant pairwise comparison in at least one of the conditions. The condition that yielded the largest number of significant genes was the higher coal tar concentration (4 μg/ml) after 8 hours of recovery. This treatment significantly affected 71 genes (47 up and 24 down). In the late time points, it was the lower concentration of coal

tar that caused the highest frequency of differential gene expression (82 genes); however, the conditions that included the higher concentration of coal tar yielded very similar differential gene expression results (75 genes). There were 56 genes in common between the high and low concentrations for the late time points.

Table 3.1 Number of significant pairwise comparisons broken down by experimental condition.

	Early time points				Late time points			
	D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
negative fold change	0	2	4	0	8	6	9	24
positve fold change	0	19	1	2	34	13	57	47
Total # of changed genes	0	21	5	2	42	19	66	71

Table 3.2 Significantly differentially expressed genes (FDR adjusted p<0.05) as determined by MAANOVA, for early recovery time points only. Values are fold changes relative to control.

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Nr4a1	nuclear receptor subfamily 4, group A, member 1	*2.07	*-1.44	**-5.70	**2.25	1.00	1.05	1.01	1.28
Cyp1a1	cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Nfil3	nuclear factor, interleukin 3, regulated	*1.76	1.11	**-3.10	*1.87	1.02	*1.59	-1.02	-1.08
Lrrn6a	leucine rich repeat neuronal 6A	1.08	**3.03	1.04	1.19	1.18	1.38	-1.09	-1.22
Tiparp	TCDD-inducible poly(ADP-ribose) polymerase	-1.41	**2.97	*1.56	1.05	**6.33	**6.25	**4.60	**5.40
Tnfrsf19	tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.00	**1.26	*-1.20	1.10
Slc2a6	solute carrier family 2 (facilitated glucose transporter), member 6	-1.07	**2.72	*-1.62	-1.19	**3.30	*1.87	**2.29	*1.66
Ckmt1	creatine kinase, mitochondrial 1, ubiquitous	1.17	**2.65	-1.27	-1.02	**4.03	*1.90	**3.61	**2.61
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.13	**2.45	**2.65	**4.17	**4.48
Cpox	coproporphyrinogen oxidase	*-1.43	**2.46	1.15	-1.35	**3.29	**3.65	**3.81	**3.63

* Significant p<0.05

** Significant FDR adjusted p<0.05

D1- Concentration of coal tar used for exposure = 1 μg/ml

D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

Table 3.3 Significantly differentially expressed genes (FDR adjusted $p<0.05$) as determined by MAANOVA, for late recovery time points only. Values are fold changes relative to control.

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cyp1a1	cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	cytochrome P450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Tiparp	TCDD-inducible poly(ADP-ribose) polymerase	-1.41	**2.97	*1.56	1.05	**6.33	**6.25	**4.60	**5.40
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.12	**2.45	**2.65	**4.17	**4.48
Tbc1d16	TBC1 domain family, member 16	1.11	2.57	-1.15	1.18	**3.36	*2.80	**4.40	*2.37
Ckmt1	creatine kinase, mitochondrial 1, ubiquitous	1.17	**2.65	-1.27	-1.06	**4.03	*1.91	**3.61	**2.61
Phldb2	pleckstrin homology-like domain, family B, member 2	1.44	*-2.07	-1.06	1.06	**-2.45	-1.28	*-2.53	**-4.02
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Insig1	insulin induced gene 1	-1.41	-1.46	-1.05	-1.33	*-1.83	**-3.88	*-2.05	**-2.37
Idi1	isopentenyl-diphosphate delta isomerase (Idi1), transcript variant 2,	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

3.2.2 Condition/Gene trees

Condition trees and gene trees, combined into one heat map were created using GeneSpring 7.0. The condition trees illustrate how samples cluster together based on similarities/differences in gene expression profiles across conditions. This can be used to determine if a group of genes discriminates across the experimental conditions. In this thesis, the condition trees are associated with a dendogram that runs along the top of the heat map, grouping similarly expressed samples horizontally. For example, two samples with similar expression profiles will be situated in close proximity in a condition tree. All branches on the condition trees were coloured by concentration, and the clustering by concentration and recovery time can be seen at the bottom of each condition tree. The genes were also grouped employing a gene tree, which can identify co-regulated genes according to similarity of gene expression across all conditions. For example, two genes with similar expression will be grouped in close proximity in a gene tree. Within the condition tree/gene tree heat map, blue indicates down regulation, red indicates up regulation and yellow indicates no change, relative to the reference.

Figure 3.6 illustrates that there was minimal clustering of the control doses towards the left of the heat map, indicating that the treatments had only a small affect on 22,000 genes. Figure 3.7 and 3.8 only display genes selected after adjustment for the false discovery. They illustrate that the control samples cluster away from the other two samples treated with coal tar, indicating that there was a clear treatment effect. These results indicate that the statistical analysis was effective in narrowing down a list of genes that responded to coal tar treatment. Although the 2 treatment concentrations clustered away from the control, they did not cluster away from each other, indicating that the chosen

concentrations were not sufficiently different to elicit clear gene expression clustering. The LacZ transgene mutation assay (Figure 3.5) also did not display a great difference between the two concentrations chosen for the microarray experiment, but did show a significant difference between treated and control. However, earlier in this section, it was mentioned that there were 56 genes in common between high and low concentration of the late time points. This leaves 46 genes that were not in common between the high and low dose, indicating that the two treatment concentrations were not eliciting the same gene expression response prolfile.

The gene tree revealed clustered of genes with similar expression profiles. Figure 3.8 (early time points, FDR adjusted $p<0.05$) illustrates 2 main groups of genes: *Ckmt1, Nmnat2, Psmb10, Cd248, H6pd, Cyp1a1* was up-regulated with respect to control and *Plk3, Tnfrsf12a, Myd116, Nr4A1* responded similarly with respect to the reference, but the fold change difference compared to the control was modest. An ontology analysis of these genes indicates that the groups are associated with metabolic process and apoptosis, respectively. Figure 3.9 (late time points, FDR adjusted $p<0.05$) illustrates one gene cluster, including *Cc17, Cc12, Id1, Nfkbia, Id3, Mfsd2, Idi1, Hmgcs1, Insig1* and *Il-33*, which is associated with steroidogenesis and cytokine-cytokine interaction. Other groups of genes were clustered, but not significantly associated with any readily identifiable functional category.

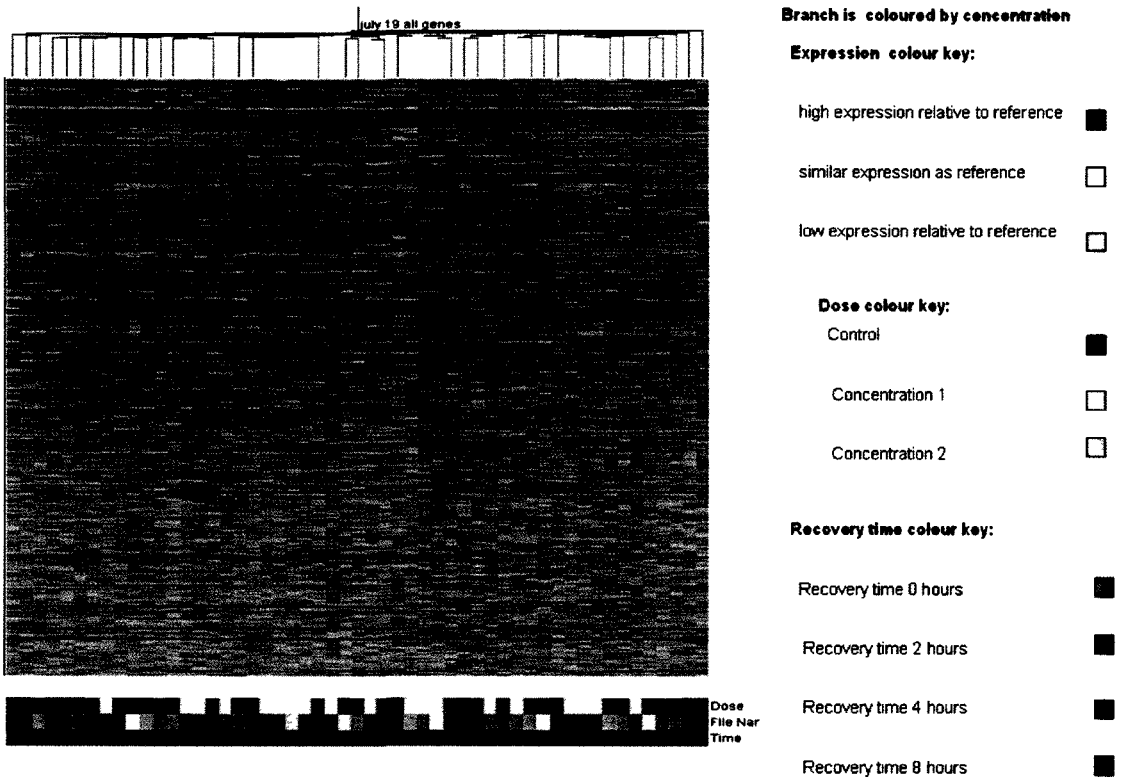


Figure 3.6 Condition tree of 22 thousand genes clustered (horizontally) according to the similarity in expression across exposure and recovery time. The branches are coloured by concentration. The colour bar below the condition tree shows how all samples cluster by concentration, file name and time. Within the condition tree heat map, blue indicates down regulation, red indicates up regulation and yellow indicates no change, relative to the reference.

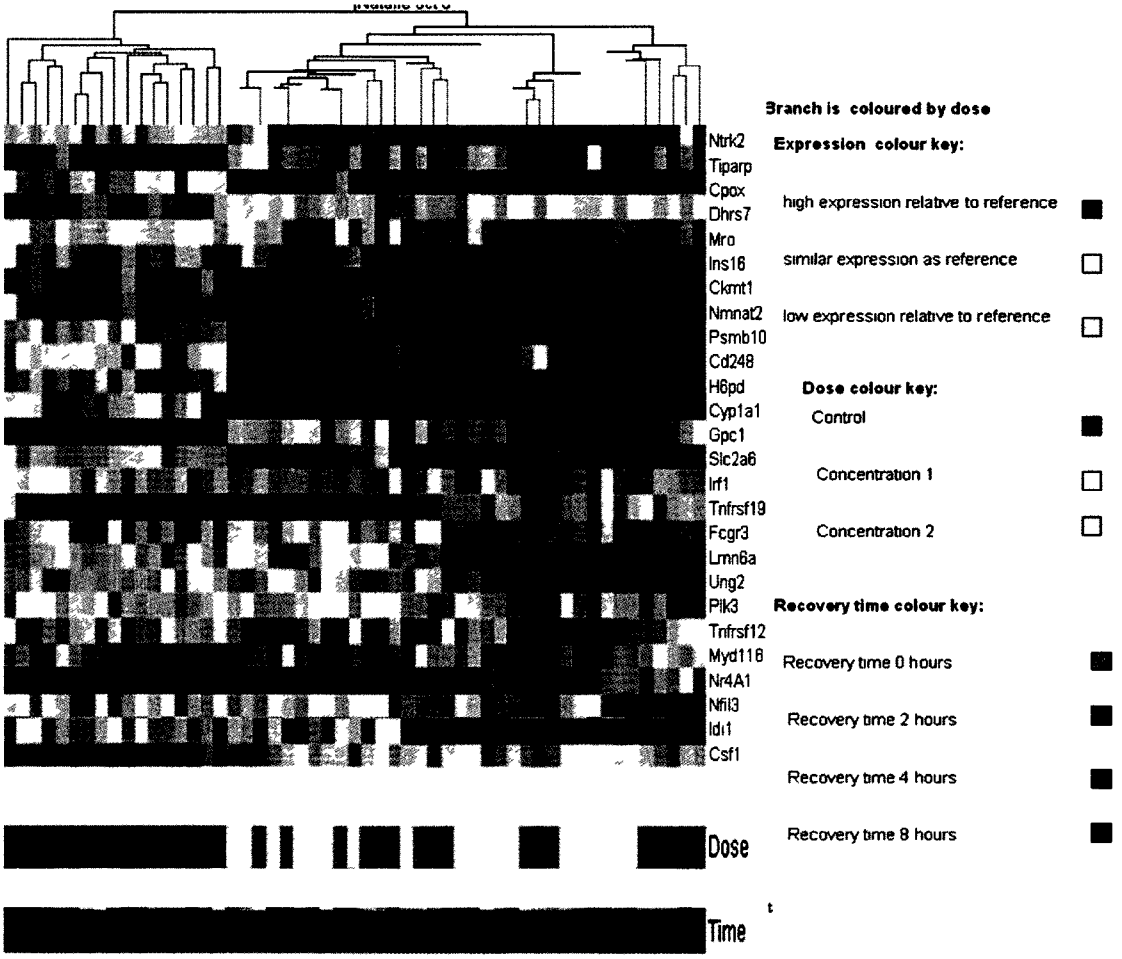
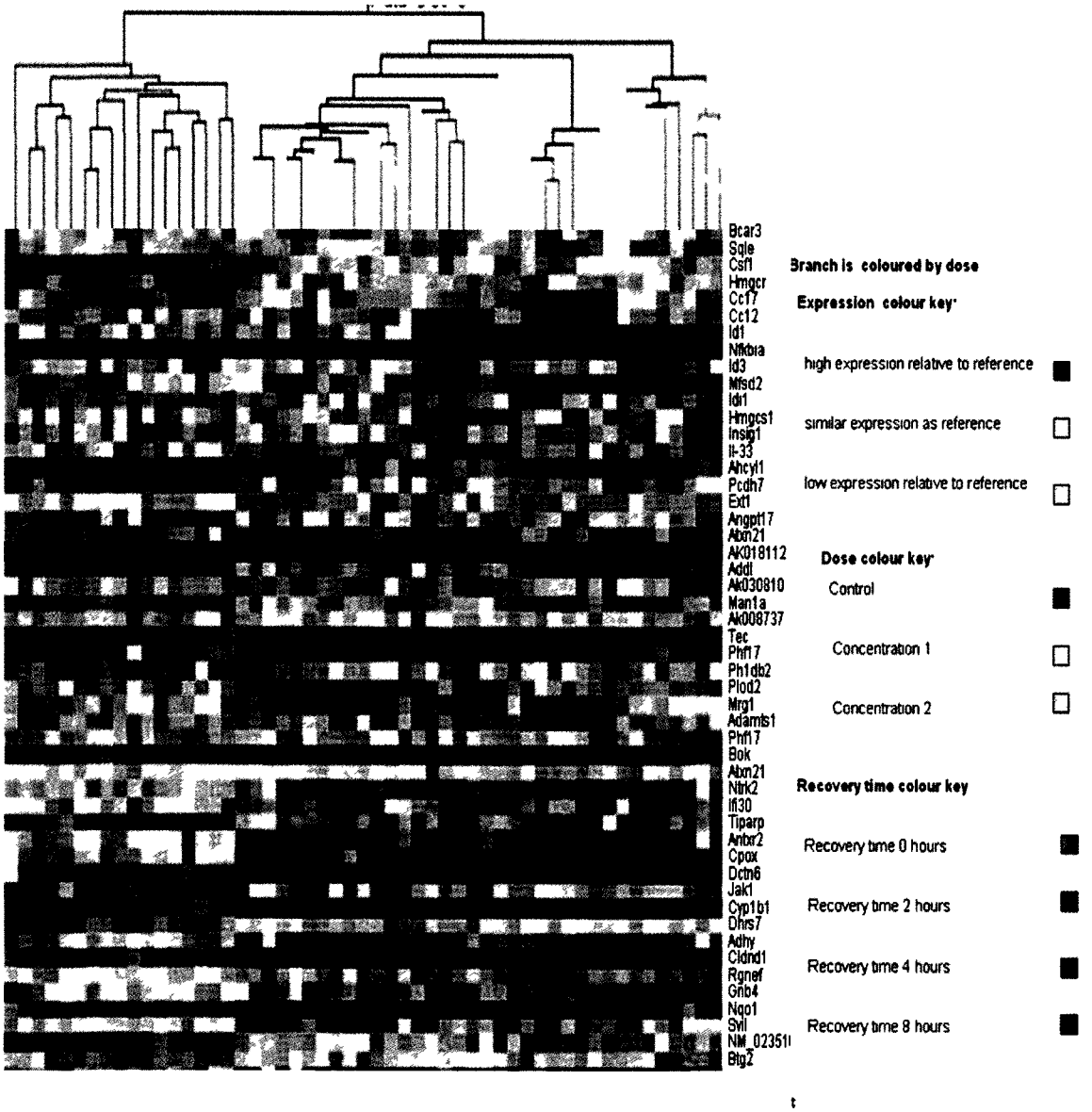


Figure 3.7 Condition tree of 26 significant genes from the early recovery time points (i.e., 0 and 2 hours recovery, FDR adjusted $p<0.05$). The branches are coloured by concentration. The colour bar below the condition tree shows how samples cluster by concentration, file name and recovery time. Within the condition tree heat map, blue indicates down regulation, red indicates up regulation and yellow indicates no change, with respect to the reference.


Bcar3
Sqle
Csf1
Hmgcr
Ccl7
Ccl2
Id1
Nfkbia
Id3
Mfsd2
Idi1
Hmgcs1
Insig1
Il-33
Ahcyl1
Pcdh7
Ext1
Angptl7
Abhd21
AK018112
Addl
Ak030810
Man1a
Ak008737
Tec
Phf17
Ph1db2
Plod2
Mrgl
Adamts1
Phf17
Bok
Abhd21
Nlrk2
Ifi30
Tiparp
Anbxr2
Cpox
Dctn6
Jak1
Cyp1b1
Dhrs7
Adhy
Clcnd1
Rgnef
Gnb4
Nqo1
Svil
NM_023511
Btg2
Branch is coloured by dose
Expression colour key
high expression relative to reference
similar expression as reference
low expression relative to reference
Dose colour key
Control
Concentration 1
Concentration 2
Recovery time colour key
Recovery time 0 hours
Recovery time 2 hours
Recovery time 4 hours
Recovery time 8 hours

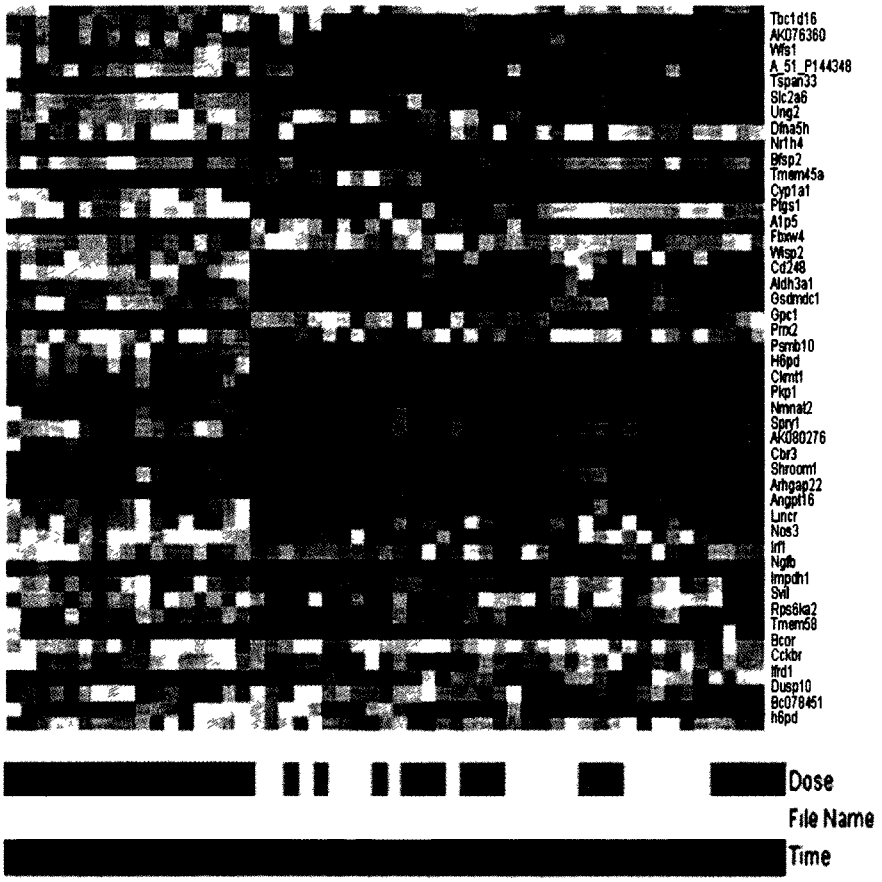


Figure 3.8 Condition tree of 102 significant genes from the later recovery time points (i.e., 4 and 8 hours recovery, FDR adjusted $p<0.05$). The branches are coloured by concentration. The colour bar below the condition tree shows how all samples cluster by concentration, file name and recovery time. Within the condition tree heat map, blue indicates down regulation, red indicates up regulation and yellow indicates no change, with respect to the reference.

3.2.3 Ontological Analyses Using David

Gene ontology analysis, which was performed in GeneSpring 7.0 and David (Database for Annotation, Visualization, and Integrated Discovery) Bioinformatics Resources, identifies gene motifs pertaining to 3 general categories: cellular components, molecular components, and biological function. The results presented in Table 3.4 show the ontology motifs, created from 457 genes that, compared to control, were (1) significantly changed with FDR adjusted $p<0.05$, (2) significantly changed with unadjusted $p<0.001$, or (3) showed a greater than 2-fold change. Table 3.5 displays the gene ontology results, created from 128 significant genes after FDR adjustment.

The data in Table 3.4 indicate that the cholesterol biosynthetic process and the pathways related to steroidogenesis are largely down regulated. Ontology categories related to angiogenesis were also mostly down regulated, but to a lesser extent than the aforementioned categories. Table 3.5 shows that the genes associated with oxidoreductase were largely up-regulated, and, once again, the genes involved in steroidogenesis displayed a down regulating trend. The genes involved in the other ontology categories in Table 3.4 and Table 3.5 do not show a clear up or down-regulation trend. Conserved between the two ontology tables are Cellular Differentiation, Blood Vessel Development (angiogenesis) and Biosynthesis of Steroids.

The percent of genes found to be differentially expressed versus the total number of genes in the pathway on the Agilent chip, was also captured in Table 3.4 and Table 3.5. The results in Table 3.4 indicate that cholesterol biosynthetic process and steroidogenesis had the highest percent of genes differentially expressed with respect to the total number of

genes in the pathway on the Agilent chip. The results in Table 3.5 indicate that ontology categories related to biosynthesis of steroids and cellular differentiation had the highest percent of genes differentially expressed with respect to the total number of genes in the pathway on the Agilent chip. Complete results of the ontological analyses can be found in Appendix D (i.e., Table 5.3 and Table 5.4).

Table 3.4 Gene ontology categories created from genes that meet one of 3 criteria: significantly changed genes after FDR adjustment ($p<0.05$), significantly changed genes with unadjusted $p<0.001$, and/or genes showing a greater than 2-fold change in expression.

Gene Ontology: Functional annotation		Number of genes in category	Percent of genes in pathway that are differentially expressed			Percent of genes in pathway that are differentially expressed – restricted those on Agilent chip	P-value		FDR
cholesterol biosynthetic process		10	0.4			43.5	0		*1.40E-08*
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Pmvk	phosphomevalonate kinase	-1.20	-1.01	-1.01	1.07	*-1.41	*-1.53	-1.11	**-2.04
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.32	**-1.85
Idi1,	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.98	*-2.16	*-1.65	**-3.86
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Nsdhl	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
Hmgsc1	3-hydroxy-3-methylglutaryl-coenzyme a synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	*-1.83	**-3.07
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	* -1.43	-1.30	**-2.12

Dhcr7	7-dehdrocholesterol reductase	1.09	-1.08	-1.38	-1.43	-1.25	*-1.13	-2.13	*-2.00
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Dhcr24	24-dehydrocholesterol reductase	-1.07	1.06	-1.02	-1.08	-1.23	-1.37	-1.08	**-1.75
biosynthesis of steroids, sterol biosynthetic process, steroid metabolic process	13	0.5			36.6		5.90E-7	*0*	
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Pmvk	phosphomevalonate kinase	-1.20	-1.01	-1.01	1.07	*-1.41	*-1.54	-1.11	**-2.04
Sqle	squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	*-1.43	**-2.65
Dhcr24	24-dehydrocholesterol reductase	1.07	1.13	1.03	1.32	-1.72	-1.40	-1.30	*-3.63
Stard4	riken cDNA 4632419c16 gene	1.19	-1.30	1.02	1.02	*-2.08	*-1.82	*-1.64	*-1.90
Dhcr7	7-dehydrocholesterol reductase	1.09	-1.08	-1.38	-1.43	-1.25	*-1.13	-2.13	*-2.00
Nqo1	nad(p)h dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36
Sc5d,	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20

Hmgcs1	3-hydroxy-3-methylglutary-coenzyme A synthase	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.08
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85
Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	*-1.43	-1.30	**-2.12
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Lss	lanosterol synthase	-1.21	-1.18	-1.03	1.17	*-1.80	*-2.05	-1.39	**-3.03
Nsdhl,	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
regulation of cell differentiation	13	0.7			7.6		4.2E-06	*0*	
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Nfkbia	nuclear factor of kappa light chain gene enhancer in b-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Tgfb2	follistatin	1.08	*-1.58	1.13	-1.13	-1.44	* -2.15	* -1.83	* -1.91
Foxg1	transforming growth	1.17	*-1.60	*-1.64	1.02	-1.30	1.06	-1.27	1.04

	factor, beta 2								
Runx1	runt related transcription factor 1	1.71	-1.10	-1.61	1.26	*1.66	*2.07	1.35	1.04
Mapk1	mitogen activated protein kinase 1	-1.01	-1.29	1.18	-1.27	*-1.33	1.08	-1.20	-1.59
Bmp4	bone morphogenetic protein 4	1.27	*-1.60	-1.47	-1.30	**-2.24	*-1.96	-1.16	*-1.56
Foxg1	forkhead box g1	1.17	*-1.60	-1.64	1.02	-1.30	1.06	-1.27	1.04
hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	*-2.49	-1.15	*-1.82	-1.60
Il6	interleukin 6	1.69	-1.15	-1.05	**2.86	-1.35	-1.31	*-1.85	*-2.03
Nrg1	neuregulin 1	1.21	1.27	-1.16	1.04	*1.60	1.30	*1.39	1.02
wwtr1	ww domain containing transcription regulator 1	-1.49	*-2.15	-1.09	-1.13	-1.08	1.29	-1.45	-1.57
Gli3	gli-kruppel family member gli3	-1.06	*-1.26	1.26	1.00	-1.20	* -1.10	1.01	-1.14
Xdh	xanthine dehydrogenase	1.04	1.09	-1.06	1.05	-1.05	1.15	**2.48	*2.19

lipid biosynthetic process	22	0.8	14.9	4.8E-06	*0*

Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Sc5d	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20

Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85
Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86
Ptgs1	prostaglandin-endoperoxide synthase 1	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	*1.99	* 2.67
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Pmvk	phosphomevalonate kinase	-1.20	-1.01	1.00	1.07	*-1.41	*-1.54	-1.11	**-2.04
Stard4	riken cDNA 4632419c16 gene	1.19	-1.30	1.02	1.02	*-2.08	*-1.82	*-1.64	*-1.90
Pcx	pyruvate carboxylase	1.22	1.21	-1.15	1.11	*1.42	1.01	*1.60	*1.30
Dhcr24	24-dehydrocholesterol reductase	1.07	1.12	1.03	1.32	-1.72	-1.40	-1.30	**-3.63
Dctn6	dynactin 6	-1.09	*1.62	1.09	-1.07	**1.95	**2.02	**2.15	*1.70
Scd2	stearoyl-coenzyme a desaturase 2	-1.16	-1.24	1.08	-1.39	*-1.65	**-2.06	1.03	**-2.20
Ptgs2	prostaglandin-endoperoxide synthase 2	*2.26	-1.53	*-3.29	1.59	*-1.65	*1.52	*-1.15	**-1.39
Pcyt1a	phosphate cytidylyltransferase 1, choline, alpha isoform	1.10	*-1.32	1.24	1.18	-1.20	1.09	-1.24	*-1.57
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	-1.43	-1.30	**-2.12

Lss	lanosterol synthase	-1.21	-1.18	-1.03	1.17	*-1.80	*-2.05	-1.39	**-3.03
Nsdhl	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
Hmgcs1	3-hydroxy-3-methylglutaryl-coenzyme a synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.08
Elovl6	elovl family member 6, elongation of long chain fatty acids (yeast)	-1.27	-1.25	1.56	-1.45	-1.31	*-1.54	*-1.51	**-2.90

blood vessel development/vascular development, angiogenesis	23	1			9.8		1.10E-09	*0*	
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	1.40	*-1.77	**-2.79	1.35	*-1.52	-1.18	-1.32	*-1.69
Ctgf	connective tissue growth factor	1.02	*-1.45	-2.11	1.01	1.11	1.01	-1.08	1.13
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.80
Tgfb2	transforming growth factor, beta 2	1.08	*-1.58	1.13	-1.13	-1.44	*-2.15	*-1.83	*-1.91
Gja1	gap junction membrane channel protein alpha 1	-1.01	-1.34	1.02	-1.01	-1.44	-1.06	*-1.78	**-2.43
Ubp1	upstream binding protein 1	-1.21	1.01	2.20	1.42	* -1.31	-1.03	1.12	1.14

Dicer1	dicer1, dcr-1 homolog (drosophila)	1.05	-1.20	1.12	1.11	-1.39	-1.08	-1.02	-1.22
Edn1	endothelin 1	1.38	*-1.47	-1.41	-1.13	*-1.73	-1.14	*-2.05	-1.42
Runx1	runt related transcription factor 1	1.71	-1.10	-1.61	1.26	*1.66	*2.07	1.35	1.04
Angptl4	angiopoietin-like 4	1.87	-1.01	-2.12	-1.17	*-1.01	-1.19	-1.13	*-1.02
Rtn4	reticulon 4	-1.16	-1.21	1.10	-1.04	*-2.18	-1.51	-1.22	-1.35
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Cited2	cbp/p300-interacting transactivator, with glu/asp-rich carboxy-terminal domain, 2	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62
Bmp4	bone morphogenetic protein 4	1.27	*-1.60	-1.47	-1.30	**-2.24	* -1.96	-1.16	*-1.57

Aw125753	riken cDNA 4731402f03 gene	1.83	-1.36	-1.11	-1.03	* -1.74	* -1.86	-1.72	** -2.66
Hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	*-2.49	-1.15	*-1.82	-1.60

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

Table 3.5 Gene ontology categories created from the list of genes that showed significant expression changes in at least one condition, after FDR adjustment. (Note: Not all conditions are shown)

Gene Ontology: Functional annotation		Number of genes in category	Percent of genes in pathway that are differentially expressed	Percent of genes in pathway that are differentially expressed – restricted those on Agilent chip	P-value	FDR
Oxidoreductase activity		*16*	*2.8*	*2.9*	*5.110E-05*	*0*
Gene symbol	**Gene Name**	**Fold change**				
		Early time points			Late time points	
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Dhrs7	dehydrogenase/reductase (sdr family) member 7	1.05	**1.74	1.12	1.02	**2.12	*2.22	**2.09	**2.64
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	*1.80
Nqo1	nad(p)h dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36
H6pd	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)	1.26	**2.02	-1.19	-1.01	**1.95	1.21	**1.94	**1.69
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.32	-1.85

Adh7	alcohol dehydrogenase 7 (class iv), mu or sigma polypeptide	-1.13	*1.38	-1.01	-1.10	*1.97	*2.08	2.99	2.36
Ptgs1	prostaglandin-endoperoxide synthase 1	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	1.99	2.67
Cpox	coproporphyrinogen oxidase	*-1.43	**2.46	1.15	-1.35	**3.29	**3.65	**3.82	**3.63
Aldh3a1	aldehyde dehydrogenase family 3, subfamily a1	1.00	*1.43	-1.28	1.298	2.03	1.41	3.08	2.24
Plod2	procollagen lysine, 2-oxoglutarate 5-dioxygenase 2	-1.19	*-1.40	1.22	-1.03	*-1.78	-1.13	-1.96	-1.74
Sqle	squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	-1.43	-2.65
Cyp1a1	cytochrome p450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	cytochrome p450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Cbr3	carbonyl reductase 3	1.20	*1.53	-1.23	1.09	**1.83	1.20	1.56	1.76
Impdh1	inosine 5'-phosphate dehydrogenase 1	1.13	1.06	-1.20	1.09	1.24	1.09	**1.56	**1.51
Dhrs7	dehydrogenase/reductase (sdr family) member 7	1.05	**1.74	1.12	1.02	**2.12	*2.23	**2.09	**2.64

Developmental process, multicellular organismal development, cellular differentiation	*32*	*5.6*		*13.2*		*5.90E-05*		*0.5*	
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Ngfb	nerve growth factor, beta	*1.97	1.43	*-1.97	1.34	*2.21	*1.46	1.41	*1.67
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	-1.85
Rgnef	rho interacting protein 2	-1.04	*1.40	1.23	1.17	*1.53	*1.48	**1.70	1.68
Btg2	b-cell translocation gene 2, anti-proliferative	-1.20	1.12	-1.35	*1.72	**1.91	*1.62	1.51	1.31
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.68	-2.31
Adrb2	adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	2.02	2.12
Prrx2	paired related homeobox 2	-1.03	*1.35	*-1.50	-1.02	*1.70	*1.84	2.89	2.44
Nr4a1	nuclear receptor subfamily 4, group a,	*2.07	*-1.45	**-5.71	**2.25	1.00	1.05	1.01	1.28

	member 1								
Fcgr3	fc receptor, IgG, low affinity iii	-1.08	**2.19	1.08	1.03	1.45	*1.76	*1.86	*1.65
Phf17	phd finger protein 17	1.25	*-1.69	1.05	-1.60	-1.19	1.10	-1.20	*-1.36
Fbxw4	f-box and wd-40 domain protein 4	1.00	1.49	-1.05	1.07	1.64	1.19	1.75	1.77
Adamts1	a disintegrin-like and metallopeptidse (reprolysin type) with thrombospondin type 1 motif, 1	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	-1.88	-2.44
Angptl6	angiopoietin-like 6	1.06	1.37	-1.11	1.20	1.28	1.38	**1.80	2.08
Myd116	myeloid differentiation primary response gene 116	-1.30	-1.23	-1.31	**2.26	1.06	*1.67	1.29	1.38
Nfkbia	nuclear factor of kappa light chain gene enhancer in b-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Ext1	exostoses (multiple) 1	-1.09	*-1.64	-1.09	-1.07	*-1.67	*-1.47	*-1.50	**-1.91
Shroom1	riken cDNA 1300007I22 gene	1.19	*1.54	*-1.66	1.32	**2.68	*1.75	1.76	1.54
Wisp2	wnt1 inducible signaling pathway protein 2	1.09	*1.88	-1.46	-1.01	**2.70	1.45	3.22	2.27
Bok	bcl-2-related ovarian killer protein	-1.05	-1.03	-1.16	1.10	*-1.40	-1.29	-1.94	-1.82
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	* 1.80
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.12	**2.45	**2.65	**4.17	**4.48
Ptgs1	prostaglandin-endoperoxide synthase 1	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	1.99	2.67
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Id3	inhibitor of DNA binding 3	*-1.51	-1.08	*1.74	1.15	*-1.99	**-2.51	-1.21	-1.48

Cited2	cbp/p300-interacting transactivator, with glu/asp-rich carboxy-terminal domain, 2	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62
Spry1	sprouty homolog 1 (drosophila)	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	2.30	1.81
Rps6ka2	ribosomal protein s6 kinase, related sequence 1	-1.09	1.05	-1.29	-1.26	**1.97	1.14	1.81	1.56
Dfna5h	deafness, autosomal dominant 5 homolog (human)	1.53	1.54	-1.31	1.03	**2.34	1.42	**3.38	3.32
Ifrd1	interferon-related developmental regulator 1	-1.15	-1.01	1.27	*1.85	1.08	**2.89	-1.01	1.25
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Arhgap22	rho GTPase activating protein 22	-1.13	*1.35	-1.13	1.11	1.27	*1.65	1.94	2.27
Blood vessel development, vascular development, angiogenesis	*8*	*1.4*			*3.4*		*1.7E-04*	*0.3*	
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
	tumor necrosis factor receptor superfamily, member 12a	*1.40	*-1.77	**-2.79	1.35	-1.51	-1.18	-1.32	*-1.69
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	* 1.80
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71

Adamts1	a disintegrin-like and metallopeptidse (reprolysin type) with thrombospondin type 1 motif, 1	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	-1.88	-2.44
Angptl6	angiopoietin-like 6	1.06	1.37	-1.11	1.20	1.28	1.38	1.79	2.08
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Arhgap22	rho GTPase activating protein 22	-1.13	*1.35	-1.13	1.11	1.27	*1.65	1.94	2.27
Cited2	cbp/p300-interacting transactivator, with glu/asp-rich carboxy-terminal domain, 2	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62
Biosynthesis of steroids	*5*	*0.9*			*12.8*		*2.30E-04*	*0.4*	
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Sqle	squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	-1.42	-2.65
Nqo1	nad(p)h dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36

Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	-1.85
Hmgcs1	3-hydroxy-3-methylglutary-coenzyme A synthase	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.08
Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)
x.xx up-regulation
x.xx down-regulation

3.3 KEGG Pathway Analyses

Cellular pathways affected by *in vitro* exposure to coal tar extract were identified using the KEGG (Kyoto Encyclopedia of Genes and Genomes) Pathway Database. Table 3.6 was created from 457 genes that, compared to control, were (1) significantly changed with FDR adjusted $p<0.05$ (2) significantly changed with unadjusted $p<0.001$, or (3) showed a greater than 2-fold change. Table 3.7 was created from 128 genes that were significant after FDR adjustment. The tables show that the identified pathways for the large and small gene lists are similar.

In the first table (Table 3.6), the KEGG analysis identified cancer as having the most genes associated with the pathway in. However the pairwise comparisons do not show a clear direction (up or down regulation) of expression for this pathway. Significant pathways that show a down regulation of gene expression include cytokine-cytokine receptor interaction, chemokine signaling, and steroid biosynthesis. Metabolism of xenobiotics by cytochrome p450 was the only pathway in this analysis that illustrates a pattern of up regulation relative to the control.

In the second table (Table 3.7), which employed the most statistically stringent gene list, the KEGG analysis identified cytokine-cytokine receptor pathway as having the most differentially expressed genes. As with Table 3.7, the gene expression showed a down regulation trend relative to the control. The other pathways had too few genes associated with them to make reliable observations regarding the direction of gene expression changes.

Table 3.6 Results of KEGG Pathway analyses employing genes that meet one of the following criteria (1) significantly changed genes with FDR adjusted p<0.05, (2) significantly changed genes with unadjusted p<0.001, or (3) genes showing a greater than 2-fold change.

KEGG Pathway	***Number of differentially expressed genes in category***	***Number of genes in pathway***	***Pathway identifier***
Cancer	*19*	*345*	*mmu05200*

Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Bmp4	bone morphogenetic protein 4	1.27	-1.60	-1.47	-1.30	-2.24	-1.96	-1.16	-1.56
Cdk6	cyclin-dependent kinase 6	-1.10	-1.02	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Fgf18	fibroblast growth factor 18	1.54	-1.10	-1.12	-1.11	-1.46	-2.21	-1.40	-2.70
Fgf7	fibroblast growth factor 7	1.82	1.00	-2.08	1.02	-1.63	-1.10	-1.20	1.00
Fzd4	frizzled homolog 4	1.14	-1.13	1.06	1.06	-1.04	-1.44	-2.33	-2.41
Gli3	GLI-Kruppel family member	-1.26	-2.03	1.31	1.35	-1.20	-1.10	1.01	-1.14
Hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	-2.49	-1.15	-1.82	-1.60
Il6	interleukin 6	1.69	-1.15	-1.05	2.86	-1.35	-1.31	-1.85	-2.03
Jak1	Janus kinase 1	-1.25	1.20	-1.05	-1.21	*1.97	2.27	2.08	1.75

Kitl	kit ligand	-1.09	-1.18	1.24	1.08	1.15	2.00	1.13	1.10
Mapk1	mitogen-activated protein kinase 1	-1.40	-1.40	1.12	-1.01	-1.33	1.08	-1.20	-1.59
Myc	myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Nos3	nitric oxide synthase 3	1.12	1.22	-1.13	1.14	1.29	*1.85	**	*
Ptgs2	prostaglandin-endoperoxide synthase 2	2.26	-1.53	-3.29	1.59	-1.65	1.52	2.29	1.80
Ralbp1	ralA binding protein 1	-1.27	-1.62	-1.07	-2.02	1.29	1.20	-1.15	-1.39
Runx1	runt related transcription factor 1	1.71	-1.10	-1.61	1.26	1.66	2.07	-1.14	1.01
Tgfb2	transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	1.35	1.04
Traf6	TNF receptor-associated factor 6	1.01	-1.03	1.09	1.08	-1.04	-1.12	-1.83	-1.91
MapK signalling pathway	*14*	*286*	*Mmu04010*						
Gene symbol	**Gene Name**	**Fold change**							

		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ddit3	DNA-damage inducible transcript 3	-1.63	-1.07	1.50	1.58	1.29	2.15	1.22	1.38
Dusp10	dual specificity phosphatase 10	1.22	-1.11	*-2.13	*2.01	1.43	**2.18	1.73	1.84
Dusp4	dual specificity phosphatase 4	2.39	1.04	-2.91	1.24	1.18	1.23	1.15	-1.13
Dusp6	dual specificity phosphatase 6	1.51	-1.41	-3.42	-1.49	-1.10	-1.11	-1.19	-1.20
Fgf18	fibroblast growth factor 18	1.54	-1.10	-1.12	-1.11	-1.46	-2.21	-1.40	-2.70
Fgf7	fibroblast growth factor 7	1.82	1.00	-2.08	1.02	-1.63	-1.10	-1.20	1.00
Gadd45a	growth arrest and DNA-damage-inducible 45 alpha	-1.83	-1.10	1.65	1.19	1.32	2.11	1.30	1.61
Mapk1	Mitogen-activated protein kinase receptor	-1.17	-1.40	1.12	1.00	-1.33	1.08	-1.20	-1.59
Myc	Myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	*2.07	*-1.45	**-5.71	**2.25	1.00	1.05	1.01	1.28
Ntrk2	Neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Rps6ka2	Ribosomal protein S6 kinase,	-1.09	1.05	-1.29	-1.26	**1.97	1.14	1.81	1.56

	polypeptide 2								
Tgfb2	Transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	-1.83	-1.91
Traf6	TNF receptor-associated factor 6	1.01	-1.03	1.09	1.08	-1.04	-1.12	1.09	1.26
Cytokine-cytokine receptor interaction	*12*	*272*	*Mmu04060*						
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cc12	chemokine (C-C motif) ligand 2	2.39	-1.81	-2.25	-1.04	-2.49	-2.32	-1.49	-2.46
Cc17	chemokine (C-C motif) ligand 7	3.22	-2.18	-3.07	-1.11	-2.07	-2.81	-3.00	-3.36
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	-1.68	-2.31
Csf2rb2	colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)	1.05	1.21	-1.11	1.10	-1.35	-2.05	1.12	1.25
Cxcl16	chemokine (C-X-C motif) ligand 16	2.04	2.04	-1.21	-1.10	1.24	1.31	1.15	1.05
Il6	Interleukin 6	1.69	-1.15	-1.05	2.86	-1.35	-1.31	-1.85	-2.03
Inhbb	Inhibin beta-B	-1.07	-1.23	-1.30	-1.18	1.51	1.10	1.94	2.50

Kitl	Kit ligand	-1.09	-1.18	1.24	1.08	1.15	2.00	1.13	1.10
Tgfb2	Transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	-1.83	-1.91
Tnfrsf12a	Tumor necrosis factor receptor superfamily, member12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Tnfrsf19	Tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.01	**1.26	*-1.20	1.10
Tslp	Thymic stromal lymphopoietin	1.44	1.41	1.00	1.03	1.20	1.67	1.37	1.34
Chemokine signalling	*10*	*203*	*mmu04062*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adcy1	Adenyate cyclase 1	-1.03	1.16	-1.41	-1.12	1.23	2.10	1.33	-1.10
Ccl2	Chemokine (C-C motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	-1.49	-2.46
Ccl20	Chemokine (C-C motif) ligand 20	1.20	1.14	-1.44	1.48	1.66	2.32	1.79	1.22
Ccl7	Chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Cxcl16	Chemokine (C-X-C motif) ligand 16	2.04	2.04	-1.21	-1.10	1.24	1.31	1.15	1.05
Gnb4	guanine nucleotide binding protein (G protein), beta 4	1.29	*2.03	-1.11	-1.18	*1.88	**2.85	2.58	2.25
Mapk1	Mitogen-activated protein kinase 1	-1.17	-1.40	1.12	1.00	-1.33	1.08	-1.20	-1.59
Nfkbia	Nuclear factor of	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60

	kappa light polypeptide gene enhancer in B-cells inhibitor-alpha								
Rock2	Rho-associated coiled-coil containing protein kinase 2	-1.08	-1.37	1.29	1.19	-1.48	-1.15	-1.81	-1.50
Shc4	Src homology 2 domain containing family, member 4	1.00	1.05	1.14	1.09	-1.20	-1.18	1.25	1.14
Jak-STAT signaling	*7*	*160*	*Mmu 4630*						
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Csf2rb2	colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)	1.05	1.21	-1.11	1.10	-1.35	-2.05	1.12	1.25
Il6	Interleukin 6	1.69	-1.15	-1.05	2.86	-1.35	-1.31	-1.85	-2.03
Jak1	Janus kinase 1	-1.25	1.20	-1.05	-1.21	*1.97	**2.27	2.08	1.75
Myc	Myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Socs3	Suppressor of cytokine signalling 3	2.06	1.30	-1.17	1.08	1.17	-1.03	1.29	1.18
Spry1	Sprouty homolog 1	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	2.30	1.81
Tslp	Thymic stromal lymphopoietin	1.44	1.41	1.00	1.03	1.20	1.67	1.37	1.34
Steroid biosynthesis	*7*	*17*	*Mmu 00100*						
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8

Cyp51	Cytochrome P450, family 51	-1.45	-1.45	1.08	1.10	-1.33	-1.43	-1.30	-2.12
Dhcr24	24-dehydrocholesterol reductase	-1.07	1.06	-1.02	-1.08	-1.23	-1.37	-1.08	-1.75
Hsd17b7	Hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	-1.96	-1.72	-1.66	-4.47
Lss	Lanosterol synthase	-1.21	-1.18	-1.03	1.17	-1.80	-2.05	-1.39	-3.03
Nsdh1	NAD(P) dependent steroid dehydrogenase 7	1.17	-1.14	-1.09	1.01	-1.07	-1.01	1.03	1.05
Sc5d	Sterol-C5-desturase	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20
Sqle	Squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	-1.43	-2.65
Metabolism of xenobiotics by cytochrome P450	5	30	Mmu0980						
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adh7	Alcohol dehydrogenase 7 (class IV)	-1.13	*1.38	-1.01	-1.10	*1.97	*2.08	2.99	2.36
Aldh3a1	Aldehyde dehydrogenase family 3, subfamily A1	1.00	*1.43	-1.28	1.30	2.03	1.41	3.08	2.24
Cyp1a1	Cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	Cytochrome P450, family 1, subfamily b,	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92

	polypeptide 1								
Gstm6	Glutathione S-transferase, mu 6	1.76	1.39	1.26	1.38	-1.66	-1.23	-1.22	-1.70

* significant $p<0.05$

** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml

D2- Concentration of coal tar used for exposure = 4 μg/ml

Tx- Hours of recovery time

(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation

x.xx down-regulation

Table 3.7 Results of KEGG Pathway analyses created from the FDR adjusted p<0.05 gene list.

KEGG Pathway	***Number of genes in category***	***Pathway identifier***
Cytokine-cytokine receptor interaction	*5*	*Mmu04060*

Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cc12	chemokine (C-C motif) ligand 2	2.39	-1.81	-2.25	-1.04	**-2.49	**-2.32	*-1.49	**-2.46
Cc17	chemokine (C-C motif) ligand 7	3.22	-2.18	-3.07	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.67	-2.31
Tnfrsf12a	Tumor necrosis factor receptor superfamily, member12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Tnfrsf19	Tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.01	**1.26	*-1.20	1.1

cancer	*4*	*Mmu05200*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cdk6	cyclin-dependent kinase 6	-1.1	*-1.56	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Jak1	Janus kinase 1	-1.25	1.2	-1.05	-1.21	*1.97	**2.27	**2.08	*1.75
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Nos3	nitric oxide synthase 3	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
MapK signalling pathway	*4*	*Mmu04010*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Dusp10	dual specificity phosphatase 10	1.22	-1.11	*-2.13	*2.01	1.43	**2.18	*1.73	*1.84

Nr4a1	Nuclear receptor subfamily 4, group A, member 1	*2.07	*-1.45	**-5.70	**2.25	1	1.05	1.01	1.28
Ntrk2	Neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Rps6ka2	Ribosomal protein S6 kinase, polypeptide 2	-1.09	1.05	-1.29	-1.26	**1.97	1.14	**1.81	*1.56
Chemokine signalling	*4*	*mmu04062*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ccl2	Chemokine (C-C motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	*-1.49	**-2.46
Ccl7	Chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Gnb4	guanine nucleotide binding protein (G protein), beta 4	1.29	*2.03	-1.11	-1.18	*1.88	**2.85	**2.58	**2.25

Metabolism of xenobiotics by cytochrome P450	*4*	*Mmu0980*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adh7	Alcohol dehydrogenase 7 (class IV)	-1.13	*1.38	-1.01	-1.1	*1.97	*2.08	**2.99	*2.36
Aldh3a1	Aldehyde dehydrogenase family 3, subfamily A1	1	*1.43	-1.28	1.298	2.03	1.41	**3.08	**2.24
Cyp1a1	Cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	Cytochrome P450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Neurotrophin signalling	*3*	*Mmu04722*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Ntrk2	Neurotrophic tyrosine kinase	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49

Rps6ka2	Ribosomal protein s6 kinase, polypeptide 2	-1.09	1.05	-1.29	-1.26	**1.97	1.14	**1.81	*1.56
Small cell lung cancer	*3*	*Mmu05222*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cdk6	Cyclin-dependent kinase 6	-1.1	*-1.56	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Nfkbia	Nuclear factor of kappa light polypeptide geneenhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Nos3	Nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
Calcium signalling	*3*	*Mmu4020*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adrb2	Adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	*2.02	**2.12

Cckbr	Cholecystokinin B receptor	**-1.05	-1.06	-1.07	-1.05	**2.18	*1.50	*1.44	1.11
Nos3	Nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
Nod-like receptor signalling	*3*	*Mmu4621*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ccl2	Chemokine (C-C motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	*-1.49	**-2.46
Ccl7	Chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Nfkbia	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Terpenoid backbone biosynthesis	*3*	*Mmu0380*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85

Hmgcs1	3-hydroxy-3-methylglutary-coenzyme A synthase	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	*-1.83	**-3.07
Idil	Isopentenyl-diphosphate delta isomerase	-1.57	-1.57	1.58	-1.36	-1.97	-2.15	-1.65	-3.86
Jak-stat	*2*	*Mmu04630*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Jak1	Janus kinase 1	-1.25	1.2	-1.05	-1.21	*1.97	**2.27	**2.08	*1.75
Spry1	Sprouty homolog 1	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	*2.30	*1.81

TGF-beta signalling	*2*	*Mmu04350*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Id1	Inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Id3	Inhibitor of DNA binding 3	*-1.51	-1.08	*1.74	1.15	*-1.99	**-2.51	-1.21	-1.48

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

3.4 Secreted proteins

In order to identify possible biomarker candidates that are secreted into the extracellular matrix, DAVID ontology analysis and Signal P (http://www.cbs.dtu.dk/services/SignalP, Technical University of Denmark, Lyngby, Denmark) were used to identify genes that are secretome candidates. Seventeen genes were found to have signal peptides indicating that they likely code for proteins that are secreted. Of these, the genes that showed the largest expression change, relative to the control, are *Ccl7, Ccl2, Csf1, Antxr2, Ptgs2*, *Gpc1* and *Dhrs7*. These results are summarized below in Table 3.8. The MAANOVA indicated that these genes were affected by coal tar concentration. None of these genes showed significance for the dose-time interaction main effect, indicating the concentration of exposure is independent of the recovery time. In other words, the effects of coal tar are not short-lived or transient, making these genes good biomarker candidates. The ontology and pathway analyzes indicates that *Cc17, Cc12* and *Csf1* are involved in cytokine-cytokine interaction, *Csf1*, *Ptgs2, Gpc1* are involved in cell differentiation and cancer and *Dhcr7* is involved in steroidogenesis.

Table 3.8 Statistically significant genes (i.e., FDR adjusted p<0.05) that code for secreted proteins.

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ccl2	chemokine (c-c motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	-1.50	-2.46
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.68	-2.31
Angptl6	angiopoietin-like 6	1.06	1.37	-1.11	1.20	1.28	1.38	**1.80	2.07
Adamts1	a disintegrin-like and metallopeptidse (reprolysin type) with thrombospondin type 1 motif, 1	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	-1.88	-2.44
Wisp2	wnt1 inducible signaling pathway protein 2	1.09	*1.88	-1.46	-1.01	**2.70	1.45	3.22	2.27
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.12	**2.45	**2.65	**4.17	**4.48
Tnfrsf19	tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.01	**1.26	*-1.20	1.10
Insl6	insulin-like 6	1.11	**1.90	1.28	1.13	**2.12	*1.64	1.24	**1.52
H6pd	hexose-6-phosphate dehydrogenase	1.26	**2.02	-1.19	-1.01	**1.95	1.21	**1.94	**1.69
Fcgr3	Fc receptor, IgG, low affinity III	-1.08	**2.18	1.09	1.03	1.45	*1.76	*1.86	*1.65
Cd248	CD248 antigen, endosialin	1.01	**2.37	*-1.81	-1.03	**2.47	*1.73	**2.98	*2.28

Plod2	procollagen lysine, 2-oxoglutarate 5-dioxygenase 2	-1.19	*-1.40	1.22	-1.03	*-1.78	-1.13	-1.96	-1.74
Ntkr2	neurotrophic tyrosine kinase, receptor, type 2, transcript variant 1	1.36	2.27	1.24	1.26	-1.07	1.21	1.09	1.12
Ptgs2	prostaglandin-endoperoxide synthase 1	2.26	-1.53	-3.29	1.59	2.14	1.64	1.99	2.67
Antxr2	anthrax toxin receptor 2	-1.29	*1.62	1.19	-1.34	**3.12	*2.49	2.93	3.03
Dhrs7	dehydrogenase/reductase (SDR family) member 7	1.05	**1.74	1.12	1.02	**2.12	*2.23	**2.09	**2.64
Cc17	chemokine (c-c motif) ligand 7	3.22	-2.18	-3.07	-1.11	-2.07	-2.81	-3.00	-3.36

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

Chapter 4: Discussion

4 Discussion

The purpose of this study was to elucidate the mechanisms of action of coal tar extract on cultured murine cells using DNA microarrays and accompanying ontological and pathway analysis; and moreover, to identify candidate biomarkers, based on differential gene expression. Two coal tar concentrations (i.e., 1 and 4 μg/ml) were selected based on the results of several cell viability, cytoxicity, and mutagenicity assays. The results of these assays indicated that there was at least 60% cell viability and an almost 2-fold increase in mutant frequency at the *lacZ* transgene at both coal tar concentrations.

The over all effect of coal tar exposure on gene expression in FE1 murine pulmonary epithelial cells is illustrated in Figure 3.6. This condition tree indicates that there is some clustering on both concentration and time. When the statistical analysis was applied to time-specific microarray data, the samples clustered more consistently by concentration. Figure 3.7 and 3.8 the clustering of samples for the early and late time points, respectively, when only statistically significant genes (i.e., false discovery rate adjustment of 0.05) are considered. The exposure had a relatively small effect on gene expression when cells were examined at the early time points (i.e., 0 and 2 hours recovery after exposure). For example, only 26 out of approximately 22,000 genes represented on the DNA microarray were found to be statistically different from the control. This was likely due to a switched sample in one of the early time point samples. There were many more statistically significant genes in the later time points, therefore, with respect to the identification of robust gene expression changes elicited by coal tar extract, the genes found to be differentially expressed at the later time points (i.e., 4 and 8 hours after exposure) are likely more useful.

Not unexpectedly, the metabolism of xenobiotics was found to be a significant gene expression motif resulting from coal tar exposure. The genes found to be involved in this pathway are all associated with responses mediated by the aryl hydrocarbon receptor, and include Cytochrome P450, family 1, members 1a (*Cyp1a1*) and b1 (*Cyp1b1*) (Vondracek *et al.* 2009), Alcohol Dehydrogenase 7 (*Adh7*) and Aldehyde Dehydrogenase family 3, subfamily A 1 (*Asdh3a1*) (Hanlon *et al.* 2005). The aryl hydrocarbon receptor (AhR) is a ligand-activated transcription factor, well known as a polycyclic aromatic hydrocarbon receptor. It plays an important role in mammalian responses to polycyclic environmental pollutants including PAHs, polychlorinated dibenzodioxins, polychlorinated biphenyls (PCBs) and polychlorinated dibenzofurans. AhR is normally located within the cytoplasm in a complex with its chaperone proteins, Hsp90, hepatitis B virus X-associated protein (XAP2) and p23. AhR-ligand complexes containing the AhR nuclear translocator protein (ARNT) ultimately bind to xenobiotic response elements (XREs) and up-regulate the transcription of target genes such as *Cyp1a1, Cyp1b1, Adh7 and Asdh3a1* (Figure 4.1). AhR has a protective role in mammals via the activation of phase I detoxification genes like *Cyp1a1* and *Cyp1b1*, which ultimately converts xenobiotics into more soluble metabolites that can be conjugated to endogenous compounds and excreted via the bile. It also plays an important role in transforming carcinogenic PAHs into metabolites that readily interact with DNA (e.g., diol epoxides). Figure 4.2 provides a schematic of the AhR signaling pathway.

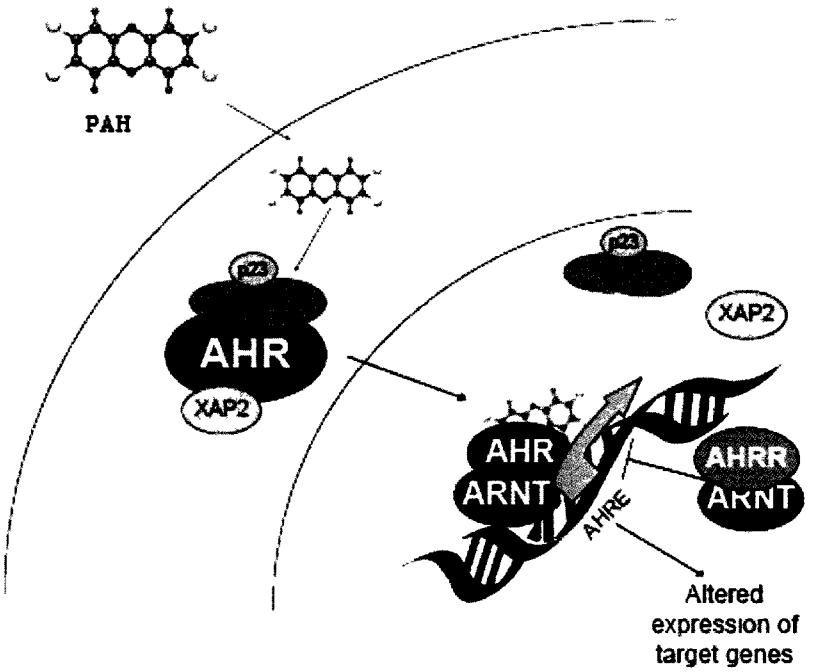


Figure 4.1 Schematic diagram of the AHR signaling pathway. AHR is a cytoplasmic receptor that, in the absence of a ligand, is complexed to several chaperone proteins. When a PAH binds to the AHR receptor, the complex is translocated to the nucleus. Here it dissociates from the chaperone proteins and binds with ARNT. The ligand/AhR/ARNT complex can then bind to specific DNA sequences called xenobiotic response elements (XREs) or aryl hydrocarbon response elements (AHREs) and up-regulate a series of unlinked genes (MacAulay 2008).

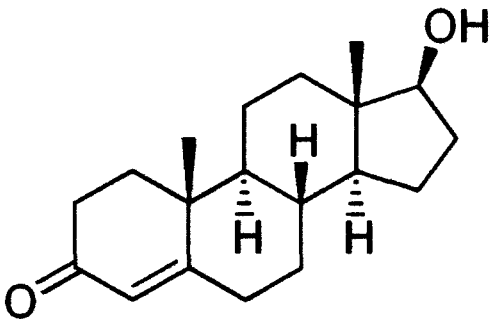
Cyp 1a1 was found to be the most highly up-regulated gene in the later time points, with the highest fold change value of 40, and *Cyp 1b1* had the second highest fold change value of 6 at concentration 1 (i.e., 1 μg/ml) for the 8 hour recovery time. A high response from the *Cyp* genes, and a positive response from other genes responsible for xenobiotic metabolism, indicates that the coal tar treatment was sufficient to elicit differential gene expression. Moreover, the microarray analysis was sensitive enough, at the concentration and recovery time points selected, to assess the differential gene expression.

The main effects of the factorial ANOVA indicate that concentration was the most significant variable that determines gene expression change. There were very few significant recovery time main effects or concentration-recovery time interactions. Time-dependent effects can confound the identification of robust biomarker candidates since they can indicate transient effects. The lack of these interactions, indicate that concentration is independant of recovery time, thereby making these genes more robust indicators of exposure. Hamadeh et al commented on the use of gene expression analysis for the identification of biomarkers, and noted that genes regulated in the same fashion following compound exposures at multiple time points can potentially serve as reliable biomarkers (Hamadeh et al. 2002b). The lack of a significant time main effect indicates that the number of recovery times investigated could be reduced in follow up experiments. This reduction would permit the analysis of more concentrations, and this in turn could provide additional insight into the cellular pathways altered by exposure to coal tar. Alternatively, the amount of time points could be expanded well beyond the 8 hours of recovery, in order to determine if the gene expression patterns remain after much longer periods of time.

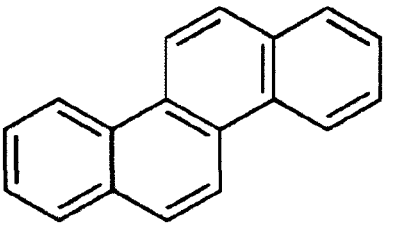
The gene expression results revealed many significant ontology categories and KEGG pathways. Though many of these pathways would be interesting to explore in depth, the remainder of the discussion will be confined to biosynthesis of steroids, angiogenesis, and cytokine-cytokine interactions.

The gene expression results presented in this thesis provide evidence that coal tar can disrupt endocrine homeostasis. The ontology analysis conducted using Genespring 7.0 and the David Bioinformatics Database, as well as the KEGG Pathway analysis, all indicate that steroid biosynthesis is an important pathway motif (see sections 3.2.3 and 3.3), with differential expression results indicating a suppression of steroid synthesis. While this disruption of endocrine system homeostasis is not a novel finding for coal tar, thus far it is poorly documented in the literature. A study by Buntner et al. reported that workers in a coal tar distillery, showed decreased serum cortisol levels in the most exposed group, and augmented levels of serum 17-hydroxycorticosteroid (Buntner *et al.* 1991).

Steroids are a class of lipids that are derivatives of the saturated tetracyclic hydrocarbon perhydorcyclopentanophenanthrene (Mathews 2000), and it has been suggested that the remarkable structural similarity between PAHs and steroid hormones is likely at the root of the endocrine disrupting properties of PAHs (Santodonato 1997). The drop in gene expression following coal tar exposure is likely the result of negative feedback loops, which are known to be present in steroid synthesis pathways (Okabe *et al.* 1998, Barraclough & Haller 1970, Mathews 2000), and known to occur at both transcriptional and translation levels (Mathews 2000). Figure 4.1 illustrates the structural similarity between testosterone and the PAH chrysene.



Testosterone



Chrysene

Figure 4.1 Structures of testosterone and chrysene. Like many steroids and PAHs, both compounds have multiple aromatic rings. The similarity may account for the observed disruption of genes involved in steroid metabolism.

The KEGG Pathway analysis revealed 6 differentially expressed genes that are involved in the biosynthesis of steroids, including 7-dehydrocholesterol reductase (*Dhcr7*), hydroxysteroid (17-beta) dehydrogenase (*Hsd17b7*), squalene epoxidase (*Sqle*), 2, 3-epoxysqualene-lanosterol cyclase (*Lss*) and sterol 14 alpha-demethylase (*Cyp51*). The Steroid Biosynthesis KEGG Pathway is illustrated in Figure 4.2, showing the genes that were differentially down-regulated by coal tar exposure. Other statistically significant genes (FDR adjusted $p<0.05$), which were not included in the KEGG results, but identified via gene ontology analysis, include phosphomeualonate (*Pmvk*), NAD(P)H dehydrogenase quinone 1 (*Nqo1*), 3-hydroxy-3-methylglutaryl-coenzyme-a-reductase (*Hmgcr*) and sterol-C5-desaturase (*Sc5d*). All of these genes are involved in the cholesterol biosynthesis pathway, confirming that coal tar exposure can affect the biosynthetic pathways of several steroids.

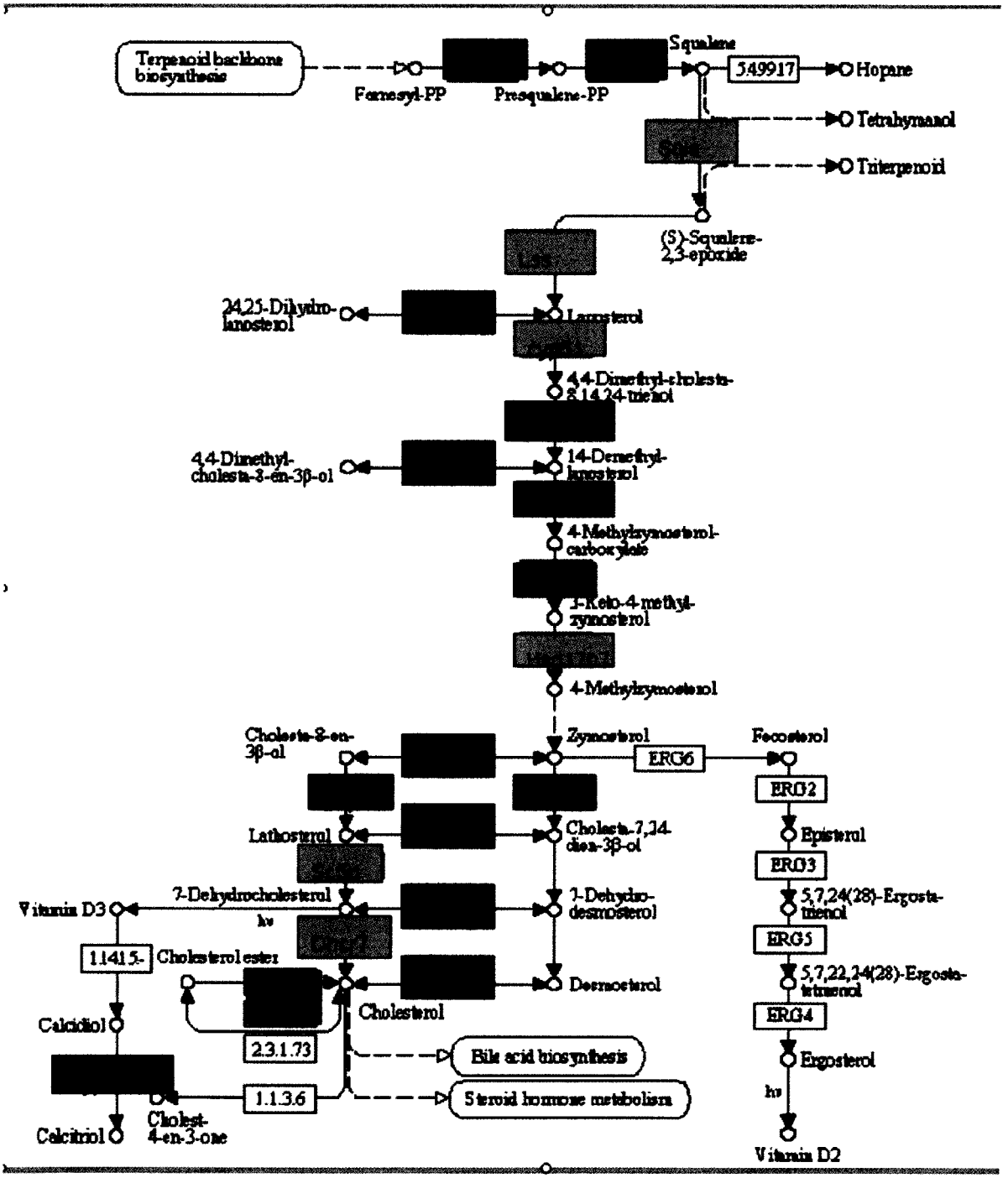


Figure 4.2 KEGG Steroid Biosynthesis Pathway. This pathway was considered significant by KEGG, David bioinformatics database and Genespring, indicating that coal tar exposure can disrupt steroid biosynthesis and endocrine homeostasis. The significantly altered genes (FDR adjusted $p<0.05$) are displayed in pink. This figure has been altered from that displayed on the KEGG Pathway website: http://www.genome.jp/KEGGbin/show_pathway?125806089516358/mmu00100.args.

Cholesterol is a precursor to all the steroid hormones, the bile acids and vitamin D. The five major classes of hormones that result from cholesterol synthesis are: progestins, which are the precursors of all other steroid hormones, glucocorticoids, which include cortisol and corticosterone that suppress inflammation reactions and promotes gluconeogenesis, minerolocorticoids, which promote ion re-absorption in the kidney, androgens (i.e., androstenedione and testosterone), which support male characteristics and sexual development, and estrogens (i.e., estrone and estradiol), which support female characteristics and sexual development. The linear C_{30} hydrocarbon called squalene (*Sqle*), whose epoxidation was found to be down regulated, is a precursor to cholesterol. Squalene is synthesized from three molecules of actyl-CoA that form 3-hydroxy-3-mehtylgultaryl-CoA (*HMG-CoA*). HMG-CoA reductase (*Hmgcr*), which was also found to be down regulated in response to coal tar extract, catalyzes an NADPH-dependent reduction of HMG-CoA to mevalonate (Mathews 2000). Hmgcr is responsible for much of the regulation of the overall steroid pathway, and the observed pattern of expression could indicate a disruption of a broad range of endocrine pathways. The steroid synthesis pathway involves a multitude of enzymes involved in oxidation and reduction of a variety of substrates, which may explain why the gene ontology analysis also revealed that the oxidoreductase motif also contains many significant genes (i.e., FDR adjusted $p<0.05$) such as *Dhrs7*, *Sqle* and *Hmgcr*.

Aside from its involvement in steroid biosynthesis, cholesterol plays a major role in membrane structure since it is one of the main lipid components of the plasma membrane in mammalian cells (Yeagle 1985). One of the occupational risks of coal tar exposure is its

photosensitizing effect, in the range of 330-550 nm. This photosensitizing effect can also be caused by progesterone, which further strengthens the argument that PAHs and endocrine hormones have a very similar mechanism of action (Choi *et al.* 2009). Moreover, the effect is also common among other HMG-CoA reductase inhibitors (i.e., a variety of statin pharmaceuticals). It has been thought that the photosensitization occurs due to an alteration in the structure and barrier properties of the skin (Feingold *et al.* 1990).

Altering the levels of cholesterol in cell membranes changes membrane fluidity (Brulet & McConnell 1976), membrane elasticity (Needham & Nunn 1990) and multiple components of the cytoskeleton (Byfield *et al.* 2004). The depletion of cholesterol has also been shown to increase the instance of apoptosis (Calleros *et al.* 2006), a cellular phenomenon that is desired and documented for the effective therapeutic use of coal tar for the treatment of skin disorders (Victor & Gottlieb 2002). Though apoptosis was not a significant ontology motif at FDR adjusted $p<0.05$, several significant genes are in the programmed cell death pathway, including tumour necrosis factor receptor superfamily member 12a (*Tnfrs12a*), DNA-damage inducible transcript 3 (Ddit3), b-cell translocation gene 2, anti-proliferative (*Btg2*), nuclear receptor subfamily 4, group a, member 1 (*Nr4a1*), PHD finger protein 17(*Phf17*), fc receptor, IgG, low affinity 3 (*Fcgr3*), bcl-2-related ovarian killer protein (*Bok*), and nitric oxide synthase (Nos3).

Related to steroidogenesis is lipid synthesis, which was also found to be down regulated in the ontology analysis summarized in Table 3.4 and Table 3.5. While the lung epithelial cells used in this study may not be expected to contribute to the inhibition of adipogenesis *in vivo*, the observed down-regulation of genes involved in this pathway is interesting since it could reveal molecular mechanisms relevant to other tissues. For

example, this finding involves some of the genes that may be involved in the weight loss effect observed after experimental exposure to BaP and cigarette smoke (Chen *et al.* 2005). The inhibition of this pathway is not surprising since both TCDD and BaP have been found to have a similar inhibitory effects on adipogenesis (Podechard *et al.* 2009). Podechard et al noted that use of an AHR antagonist (α-naphthoflavone) concomitantly with BaP, negated the inhibitory adipogenesis effect, implicating the involvement of the aryl hydrocarbon receptor (Podechard et al. 2009).

Although PAHs are remarkably similar in structure to cholesterol, there is another possible explanation for the down regulation of HMG-CoA reductase, as well as the sterol and cholesterol pathways. The PAHs in coal tar could be oxidizing endogenous sterols to form oxysterols. This reaction has been shown to be catalyzed by a variety of members of the cytochrome P450 family (Bjorkhem 2002). There is evidence that it is the oxygenated derivatives of cholesterol, rather than cholesterol itself, that causes the strongest inhibition of sterol biosynthesis by causing a complete, or almost complete, inhibition of HMG-Coa reductase activity (Kandutsch & Chen 1973) through a negative feedback loop (Schroepfer 2000). Oxysterols can be taken in through foods rich in cholesterol, or can be formed as a result of oxidative stress, and it is interesting to note that oxysterols have been identified as markers to oxidative stress (Boaz *et al.* 2005). For example, oxysterols have been found to increase in cigarette smokers, and in the skeletal muscle of rats following chronic exposure to alcohol (Adachi *et al.* 2003). Oxysterols have also been found in the lung after acute oxidative stress from paraquat (Ishii *et al.* 2002) and ozone (Pulfer & Murphy 2004). The alterations in oxidoreductase activity observed here (see results Table 3.5) could indicate that oxysterols are indeed being created from the oxidative stress caused by the coal tar exposure. Finally, the activation of MAPKs, as seen in this study as an upregulation of the

genes in this pathway, has been shown to be functionally linked to the oxidative stress response causing an inhibitory effect on steroidogenesis (Abidi *et al.* 2008).

KEGG Pathway analysis also revealed significant genes (i.e., FDR adjusted $p<0.05$) involved in cytokine-cytokine interactions, and these genes were also predominately down regulated (see results Table 3.7), indicating that coal tar elicits an immunomodulatory effect. Cytokines are composed of a large superfamily of polypeptides that play important roles in the immune system by recruiting leukocytes to sites of injury and inflammation (Strieter *et al.* 1996). The gene coding for *Csf-1*, was found to be significantly down regulated, as were *Ccl2* and *Ccl7* from the CC chemokine family. The down regulation of these cytokines is an interesting response because other complex mixtures, such as urban air particulate matter, and diesel exhaust, which also contain PAHs, have been shown to induce cytokine expression in lung epithelial cells (Ovrevik *et al.* 2009). This unique difference in gene expression may be useful in selecting a biomarker profile that is specific for coal tar exposure.

The inhibition of *Ccl2* and *Cc17* expression observed in this study has been previously associated with the anti-inflammatory activity in nephritis, arthritis, pancreatitis and colitis after treatment with an indazole derivative called bindarit (a polycyclic aromatic amine similar to many compounds in coal tar) (Zoja *et al.* 1998, Guglielmotti *et al.* 2002, Bhatia *et al.* 2005, Bhatia *et al.* 2008). The anti inflammatory action of bindarit has been attributed to the disruption of monocyte recruitment in the affected tissue (Mirolo *et al.* 2008). In a study to investigate the effect of bindarit on lipopolysaccaride induced endotoxemia, *Ccl2* (referred to as *Mcp-1* in the study) expression was effectively blocked, leading to decreased lung injury due to the decreased recruitment of neutrophils in the lungs (Ramnath *et al.* 2008). Coal tar exposure in this study appears to be causing a similar

effect on *Ccl2* expression levels, indicating that it may also be effective at blocking monocyte and neutrophil recruitment, and may have similar effects on inflamed tissues.

In another study that employed coal tar and UV light, also known as the Goeckerman regime for the treatment of psoriasis, there was a reported decrease in serum levels of *Ccl2* and other chemokines in both healthy volunteers and psoriatic patients who received the treatment. In addition, the lowered serum levels of chemokines were coupled with diminished clinical symptoms in the psoriatic patients (Pohl *et al.* 2009).

Csf-1, found to be significantly down regulated in this study, regulates recruitment, growth, survival and differentiation of macrophages. A high level of macrophage recruitment is associated with inflammatory disease (Moss & Hamilton 2000), and it has been demonstrated that an exposure to exogenous Csf-1 exasperates inflammatory conditions, whereas the inhibition of *Csf-1* reduces them (Campbell *et al.* 2000). The inhibition of cytokines in general, has been shown to occur after topical steroid treatment in allergen-induced reactions, and after *in vivo* exposure to glucocorticoids in healthy volunteers (Teoh *et al.* 1995, Sim *et al.* 1995). If PAHs are mimicking steroids, as was suggested earlier in this discussion, then the observation seen in this study that cytokine production is inhibited, is in line with the literature (i.e., that cytokine inhibition will occur after steroid exposure). The unique differential expression of cytokines found in this study, is also useful for the identification of candidate biomarkers since the primary function of cytokines is cell-cell interactions, and they are therefore excreted.

Angiogenesis was one of the significant pathways identified by the ontology analysis shown in Table 3.4. Most of the significantly altered genes in this ontology category were found to be down-regulated, suggesting that coal tar exposure elicits an anti-

angiogenic effect. Angiogenesis is the growth of new blood vessels from existing vessels. It is highly regulated, and is an important physiological process in wound healing, growth and development, and reproduction. Unregulated angiogenesis is related to diseases such as arthritis, psoriasis and cancer. In tumors, angiogenosis is necessary for malignancy and metastasis, and it has therefore been the subject of intense research. The observation that coal tar may have antiangiogenic properties is not completely surprising, since AhR agonists have previously been shown to have antiangiogenic properties (Ichihara *et al.* 2007). For example, Ichihara et al. found that the ablation of AhR in knock out mice exposed to BaP resulted in increased ischemia-induced angiogenesis (Ichihara et al. 2007), and blood flow was greatly diminished in wild-type BaP-exposed mice compared to the control (i.e. unexposed,wild-type mouse). Similar results of angiogenesis inhibition have been seen following exposures to cigarette smoke (Michaud *et al.* 2003). Ivnitski-Steele and Walker (Ivnitski-Steele & Walker 2003) demonstrated that TCDD inhibits coronary vasculogenesis, and in a later study, (Ivnitski-Steele *et al.* 2005) concluded that the TCDD inhibition of coronary vasculogenesis is regulated, in part, by the reduced responsiveness to endogenous angiogenic stimuli, including vascular endothelial growth factor A (VEGF-A). The increased expression of *Cyp1a1* has also been shown to have an antiangiogenic effect that can reduce cell proliferation in tumors, and induce developmental disruptions (Miyata *et al.* 2003). As already noted, the results obtained here show that *Cyp1a1* is by far the most highly expressed of the differentially expressed genes identified. Finally, it has been well reviewed that angiogenesis is dependant on the up regulation of *Ccl2* and *Cc17* and *Cxcr2*, the latter of which is activated by the previous two. When these cytokines are inhibited, as they are in this study, angiogenesis is also impaired (Mehrad *et al.* 2007).

As noted in the introduction, coal tar has long been known to induce tumours in mice, but the evidence to support its carcinogenicity in humans is mixed. The KEGG pathway analysis of the differentially expressed genes determined that cancer was one of the most significant motifs for coal tar exposure. On closer inspection of the significant genes in this pathway, it is difficult to draw any firm conclusions regarding the impact of coal tar exposure on genes involved in the initiation, growth and development of tumours. The lack of any decisive involvement of coal tar exposure on the cancer pathways is not completely surprising, since it has been well documented that a single PAH mixture can alter the risk of another PAH by inhibiting or augmenting their carcinogenicity (Marston *et al.* 2001).

One of the goals of this project was to identify potential biomarker candidates that could subsequently be validated using quantitative RT-PCR, protein analysis, and *in vivo* studies. Gene ontology analysis using the David Bioinformatics Database and SignalP was used to determine if any of the differentially expressed genes code for proteins that would be expected to be secreted into the extracellular matrix (i.e., culture medium *in vitro* and serum *in vivo*). Seventeen genes were found to have signal peptide sequences indicating they are likely to be secreted (see results Table 3.8). The secretome candidates that showed the largest magnitude of change with respect to control are *Ccl7, Ccl2, Csf1, Antxr2, Ptgs2, Gpc1* and *Dhrs7*. The MAANOVA indicated that these genes were affected by coal tar concentration. None of these genes were significant for the dose-time interaction main effect, indicating the concentration of exposure is independent of the recovery time. In other words, the effects of coal tar are not short-lived or transient, making these genes good biomarker. Thus, it could be argued that these are the best *in vivo* biomarker candidates

revealed in this study. However, since each of these secreted gene products alone may not be specific enough to indicate coal tar exposure, it is likely that a group of differentially expressed secreted proteins, with the magnitudes and directions of the induced changes, would be necessary for a reliable proteomic biomarker of coal tar. However, before any proteomic profile could be seriously considered as a biomarker candidate, differential expression of these secretome elements would have to be verified using proteomic techniques (i.e., Western blots, mass spectrometry). Moreover, the utility of the proteomic profile would have to be verified by subsequent analyses of protein function, protein stability, and the consistency of the response across different exposure scenarios, including *in vivo* exposures in experimental animals. Nevertheless, the successful identification of a small number of secreted proteins whose expression appears to be significantly altered by coal tar exposure underscores the utility of a combined genomic-proteomic-bioinformatic approach for the identification of biomarker candidates.

Conclusions & Future work

The exposure of coal tar to FE1 cells elicited changes in gene expression with respect to the control, and the differential gene expression provided mechanistic insights by revealing that the exposure elicited inhibitory effects on steroidogenesis, cytokine-cytokine interaction and angiogenesis. The study also reavealed 6 biomarker candidates that include *Ccl7, Ccl2, Antxr2, Ptgs2, Gpc1* and *Dhrs7.*

Definitive identification and validation of precise biomarkers for complex mixtures of environmental carcinogens such as coal tar will require several complementary research endeavours. Quantitative RT-PCR will be required to confirm the highlighted gene expression signatures, and this effort should include the genes that code for proteins identified as secretome candidates. Once this is completed, *in vivo* studies on coal tar or other complex environmental matrices (e.g., contaminated soil, urban air particulates, diesel exhaust), using the appropriate route of exposure (i.e., inhalation, dermal application, oral), would be essential to validate the list of biomarker candidates. Comparisons of gene expression and proteomic profiles across mixtures for a variety of exposure scenarios should ultimately permit the identification of biomarkers that reliably indicate exposure to, or effects from, a specific complex mixture like coal tar. This may prove challenging since complex environmental mixtures containing similar components (e.g., combustion by-products) may yield similar genomic and proteomic signatures. It will be necessary to use pattern recognition and dimensional reduction algorithms, such canonical discriminant analysis, to identify profiles that can discriminate between the different exposure scenarios. This and other types of dimension reduction techniques have been used to generate gene

expression profiles, while allowing for inter-animal differences (Hamadeh et al. 2002b). The work presented in this thesis constitutes completion of the initial stages of the work required to identify a robust biomarker of coal tar exposure in cultured cells, and eventually, in experimental animals and humans.

5 Appendices

5.1 Appendix A

Experimental design and statistical analysis of microarray data - by Andrew Williams

A reference design (Kerr & Churchill 2001, Kerr 2003) with the median signal intensities was used to analyze the Coal Tar data. The experiment included main effects of concentration (3 levels), time (4 levels) and concentration by time interactions. Five biological replicates per condition were used for a total of 60 microarrays. Two microarrays were removed from the study due to poor data quality.

The background for each array was measured using the (-) 3xSLv1 probe. Spots with median signal intensities less than the trimmed mean (trim = 5%) plus three trimmed standard deviations of the (-) 3xSLv1 probe were flagged. The total number of flagged spots, the median signal intensity and standard deviation for the (-) 3xSLv1 probe for each array were recorded. Other array level summary statistics included the median signal to noise ratio ($\log_2$ scale) for each channel. This information was used to help identify microarrays with poor data quality.

All pre-processing of the data was conducted using R (Yang *et al.* 2002). The data were normalized using the global Lowess method (RDCT 2005) using the transform.madata function in the MAANOVA library (Wu *et al.* 2003). Ratio intensity plots were constructed for the raw and normalized data for each array. Hierarchical clustering using complete and single linkages were generated. Through inspection of the dendrogram, no outliers or suspect samples were detected.

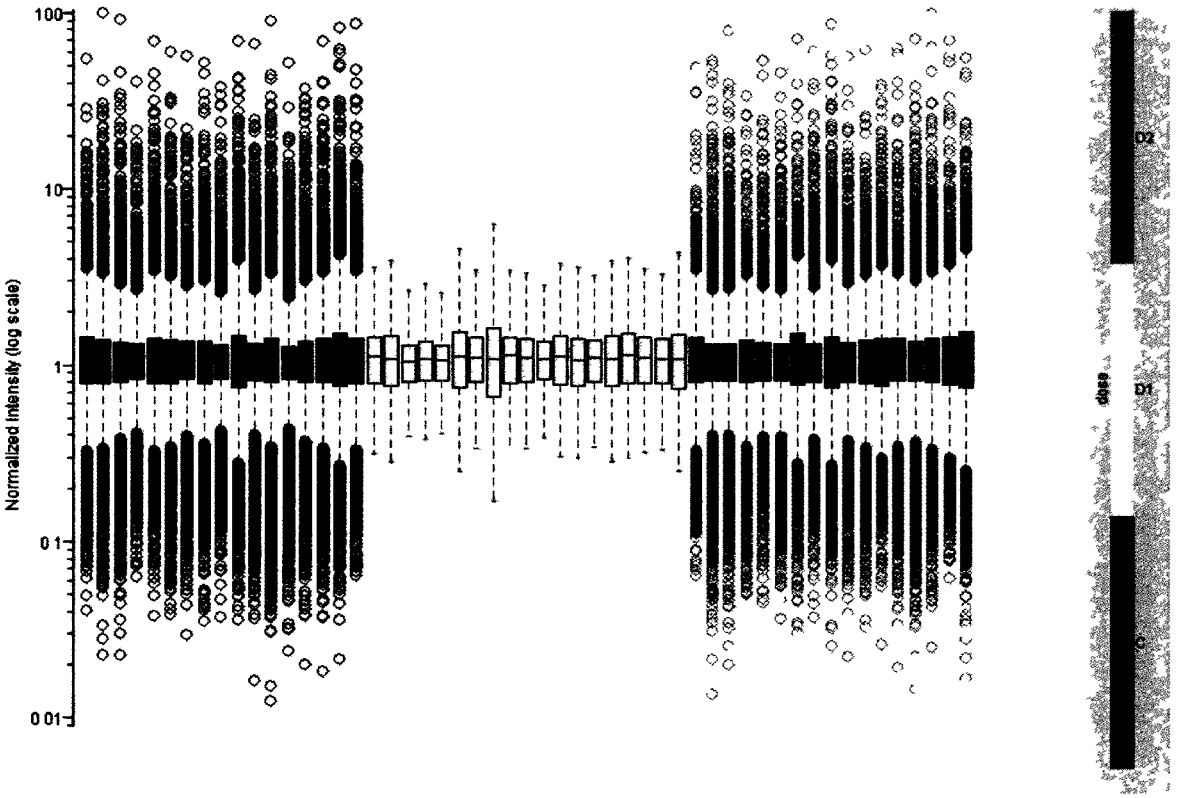
Differentially expressed genes between the control and exposed samples at the four time points were tested using the MAANOVA library. The ANOVA model was fitted to include the main effects of concentration and time, with a concentration by time interaction term. The Fs statistic (Cui *et al.* 2005), a shrinkage estimator, was used for the gene-specific variance components and the associated p-values for all the statistical tests were estimated using the permutation method (10,000 permutations with residual shuffling). These p-values were then adjusted for multiple comparisons by using the false discovery rate approach (Benjamini & Hochberg 1995).

The least-squares means(Searle 1980) a function of the model parameters, was used to estimate the fold changes for each pairwise comparison.

5.2 Appendix B

Box Plots Constructed From Normalized Microarray Data

Box plots were constructed in Genespring 7.0 (Agilent Technologies) on Lowess-normalized microarray data. Box plots are displayed below.



5.3 Appendix C

Complete list of genes found to be differentially expressed after FDR adjustment ($p<0.05$)

Table 5.1 Significantly differentially expressed genes (FDR adjusted $p<0.05$) as determined by MAANOVA results employing only early recovery time points in the statistical analysis.

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Nr4a1	nuclear receptor subfamily 4, group A, member 1	*2.07	*-1.44	**-5.70	**2.25	1.00	1.05	1.01	1.28
Cyp1a1	cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Nfil3	nuclear factor, interleukin 3, regulated	*1.76	1.11	**-3.10	*1.87	1.02	*1.59	-1.02	-1.08
Lrrn6a	leucine rich repeat neuronal 6A	1.08	**3.03	1.04	1.19	1.18	1.38	-1.09	-1.22
Tiparp	TCDD-inducible poly(ADP-ribose) polymerase	-1.41	**2.97	*1.56	1.05	**6.33	**6.25	**4.60	**5.40
Tnfrsf19	tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.00	**1.26	*-1.20	1.10
Slc2a6	solute carrier family 2 (facilitated glucose transporter), member 6	-1.07	**2.72	*-1.62	-1.19	**3.30	*1.87	**2.29	*1.66
Ckmt1	creatine kinase, mitochondrial 1, ubiquitous	1.17	**2.65	-1.27	-1.02	**4.03	*1.90	**3.61	**2.61
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.13	**2.45	**2.65	**4.17	**4.48
Cpox	coproporphyrinogen oxidase	*-1.43	**2.46	1.15	-1.35	**3.29	**3.65	**3.81	**3.63
Cd248	CD248 antigen, endosialin	1.01	**2.37	*-1.81	-1.03	**2.48	*1.73	**2.98	*2.28

Myd116	myeloid differentiation primary response gene 116	-1.30	*-1.23	-1.31	**2.26	1.06	*1.67	1.29	1.38
Plk3	polo-like kinase 3	1.18	*-1.38	**-2.26	*1.69	-1.28	-1.30	-1.29	*-1.43
Fcgr3	Fc receptor, IgG, low affinity III	-1.08	**2.18	1.08	1.03	1.45	*1.76	*1.86	*1.65
Id1	inhibitor of DNA binding 1	1.02	*-2.06	**1.44	-1.29	**-3.46	**-3.93	*-2.02	**-2.71
H6pd	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)	1.26	**2.02	-1.19	-1.01	**1.95	1.21	**1.94	**1.69
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.68	-2.31
Psmb10	proteasome (prosome, macropain) subunit, beta type 10	1.14	**1.93	1.10	1.11	**1.80	1.45	**2.16	**1.76
Insl6	insulin-like 6	1.11	**1.90	1.28	1.13	**2.12	*1.64	1.24	**1.52
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	*1.40	*-1.77	**-2.79	1.359	-1.52	-1.18	-1.32	*-1.69
Nmnat2	nicotinamide nucleotide adenylyltransferase 2	1.18	**2.23	-1.11	-1.03	**3.45	*1.88	**2.68	*1.60
Irf1	interferon regulatory factor 1	*-1.34	**1.87	**2.12	-1.18	**2.26	*1.69	**1.86	**1.81
Mro	Maestro	1.00	**1.98	1.29	1.23	*1.53	1.29	1.33	1.40
Ung2	Uracil-DNA Glycosylase 2	-1.12	**1.93	1.13	1.08	1.39	1.11	**2.16	**1.90
Dhrs7	Dehydrogenase/reductase (SDR family) member 7	1.05	**1.74	1.12	1.02	**-1.18	*1.04	**1.04	**1.06
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49

* Significant $p<0.05$

** Significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml

D2- Concentration of coal tar used for exposure = 4 μg/ml

Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

Table 5.2 Significantly differentially expressed genes (FDR adjusted $p<0.05$) as determined by MAANOVA results employing only late recovery time points in the statistical analysis

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cyp1a1	cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	cytochrome P450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Tiparp	TCDD-inducible poly(ADP-ribose) polymerase	-1.41	**2.97	*1.56	1.05	**6.33	**6.25	**4.60	**5.40
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.12	**2.45	**2.65	**4.17	**4.48
Tbc1d16	TBC1 domain family, member 16	1.11	2.57	-1.15	1.18	**3.36	*2.80	**4.40	*2.37
Ckmt1	creatine kinase, mitochondrial 1, ubiquitous	1.17	**2.65	-1.27	-1.06	**4.03	*1.91	**3.61	**2.61
Phldb2	pleckstrin homology-like domain, family B, member 2	1.44	*-2.07	-1.06	1.06	**-2.45	-1.28	*-2.53	**-4.02
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Insig1	insulin induced gene 1	-1.41	-1.46	-1.05	-1.33	*-1.83	**-3.88	*-2.05	**-2.37

Idi1	isopentenyl-diphosphate delta isomerase (Idi1), transcript variant 2,	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86
Cpox	coproporphyrinogen oxidase	*-1.43	**2.46	1.15	-1.35	**3.29	**3.65	**3.82	**3.63
Nmnat2	nicotinamide nucleotide adenylyltransferase 2	1.18	**2.23	-1.11	-1.04	**3.45	*1.88	**2.68	*1.60
Dfna5h	deafness, autosomal dominant 5 homolog (human)	1.53	1.54	-1.31	1.03	**2.34	1.42	**3.38	**3.32
Ccl7	chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Nqo1	NAD(P)H dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36
Slc2a6	solute carrier family 2 (facilitated glucose transporter), member 6	-1.07	**2.72	*-1.62	-1.19	**3.30	*1.87	**2.29	*1.66
Tmem45a	transmembrane protein 45a	1.07	*1.567	-1.23	-1.06	**2.12	1.30	**3.28	**3.14
Angptl7	angiopoietin-like 7	1.39	-1.41	1.29	-1.18	-1.17	*-2.79	*-2.21	**-3.24
Wisp2	WNT1 inducible signaling pathway protein 2	1.09	*1.88	-1.46	-1.01	**2.70	1.45	**3.22	**2.27
9230117N10 Rik	RIKEN cDNA 9230117N10 gene	-1.11	-1.28	1.22	-1.20	-1.24	*-1.67	**-2.51	**-3.18
Antxr2	anthrax toxin receptor 2	-1.29	*1.62	1.19	-1.34	**3.12	*2.49	**2.93	**3.03
Gsdmdc1	gasdermin domain containing 1	-1.06	1.38	-1.64	-1.41	**2.15	*1.89	**3.10	**2.39
Aldh3a1	aldehyde	1.00	*1.43	-1.28	1.298	2.03	1.41	**3.08	**2.24

	dehydrogenase family 3, subfamily A1								
Hmgcs1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	*-1.83	**-3.07
Pkp1	plakophilin 1	1.04	**1.78	-1.22	-1.17	**3.05	*1.83	**2.30	**1.90
Adh7	alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide	-1.13	*1.38	-1.01	-1.10	*1.97	*2.08	**2.99	*2.36
Cd248	CD248 antigen, endosialin (Cd248), mRNA [NM_054042]	1.01	**2.37	*-1.81	-1.03	**2.47	*1.73	**2.98	*2.28
9930013L23 Rik	adult male medulla oblongata cDNA, RIKEN full-length enriched library, clone:6330404C01 product:hypothetical protein, full insert sequence. [AK018112]	1.38	-1.17	1.33	1.25	*-1.82	-1.37	*-2.07	**-2.97
Prrx2	paired related homeobox 2 (Prrx2), mRNA [NM_009116]	-1.03	*1.35	*-1.50	-1.02	*1.70	*1.84	**2.89	**2.44
Ifrd1	interferon-related developmental regulator 1 (Ifrd1), mRNA [NM_013562]	-1.15	-1.01	1.27	*1.85	1.08	**2.89	1.00	1.25
Tspan33	tetraspanin 33 (Tspan33), mRNA [NM_146173]	1.02	*2.23	-1.29	-1.05	**2.86	*1.90	*2.15	**2.23

Gnb4	guanine nucleotide binding protein, beta 4 (Gnb4), mRNA [NM_013531]	1.29	*2.03	-1.11	-1.18	*1.88	**2.85	**2.58	**2.25
Ntrk2	adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500040I13 product:neurotrophic tyrosine kinase, receptor, type 2, full insert sequence. [AK018789]	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Nr1h4	nuclear receptor subfamily 1, group H, member 4 (Nr1h4), mRNA [NM_009108]	-1.02	*1.74	1.07	1.09	**2.75	*1.77	**2.48	*2.01
Shroom1	RIKEN cDNA 1300007L22 gene (1300007L22Rik), mRNA [NM_027917]	1.19	*1.54	*-1.66	1.32	**2.68	*1.75	*1.76	*1.54
Ptgs1	prostaglandin-endoperoxide synthase 1 (Ptgs1), mRNA [NM_008969]	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	*1.99	**2.67
Wfs1	Wolfram syndrome 1 homolog (human) (Wfs1), mRNA [NM_011716]	1.14	*1.54	-1.24	1.20	*1.76	*1.95	**2.65	**2.58
Sqle	squalene epoxidase (Sqle), mRNA [NM_009270]	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	*-1.43	**-2.65

Dhrs7	dehydrogenase/reductase (SDR family) member 7 (Dhrs7), mRNA [NM_025522]	1.05	**1.74	1.12	1.02	**2.12	*2.23	**2.09	**2.64
Aqp5	aquaporin 5 (Aqp5), mRNA [NM_009701]	-1.17	*1.38	-1.28	1.23	*1.85	*1.44	**2.30	**2.63
Tmem58	transmembrane protein 58 (Tmem58), mRNA [NM_175259]	1.23	-1.02	-1.49	-1.27	1.44	1.02	**2.62	1.411
AK076360	10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732466E16 product:hypothetical protein, full insert sequence. [AK076360]	1.24	1.50	-1.68	1.32	**2.60	*2.03	*1.97	*1.58
Csf1	colony stimulating factor 1 (macrophage) (Csf1), mRNA [NM_007778]	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.67	-2.31
Bcar3	breast cancer anti-estrogen resistance 3 (Bcar3), mRNA [NM_013867]	1.24	*-1.5	-1.20	-1.18	*-1.93	*-1.94	*-2.07	**-2.55
Pcdh7	protocadherin 7 (Pcdh7), mRNA [NM_018764]	*1.55	-1.42	1.11	1.06	**-2.43	*-1.90	*-1.79	**-2.51
Bfsp2	beaded filament structural protein 2, phakinin, mRNA (cDNA clone	1.10	1.21	-1.15	1.02	**2.51	1.24	**2.10	*1.45

	IMAGE:582139), complete cds. [BC062815]								
Man1a	mannosidase 1, alpha (Man1a), mRNA [NM_008548]	-1.05	-1.18	*1.36	-1.13	-1.38	*-1.48	**-2.08	**-2.51
Ahcyl1	S-adenosylhomocysteine hydrolase-like 1 (Ahcyl1), mRNA [NM_145542]	-1.14	-1.16	*1.41	-1.07	*-1.64	*-1.82	*-1.98	**-2.49
Ccl2	chemokine (C-C motif) ligand 2 (Ccl2), mRNA [NM_011333]	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	*-1.49	**-2.46
Spry1	sprouty homolog 1 (Drosophila) (Spry1), mRNA [NM_011896]	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	*2.30	*1.81
Adamts1	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 1 (Adamts1), mRNA [NM_009621]	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	*-1.88	**-2.44
Adrb2	3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630019C20 product:BETA-2 ADRENERGIC RECEPTOR, full insert sequence.	-1.17	*1.54	1.24	-1.13	**2.44	1.45	*2.02	**2.12

	[AK080276]								
Cldnd1	claudin domain containing 1 (Cldnd1), mRNA [NM_171826]	1.06	1.48	-1.28	-1.17	*2.34	*2.36	**2.17	*2.05
Ifi30	interferon gamma inducible protein 30 (Ifi30), mRNA [NM_023065]	1.10	2.02	-1.19	-1.05	*1.63	*1.64	**2.33	**2.08
Nos3	nitric oxide synthase 3, endothelial cell (Nos3), mRNA [NM_008713]	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
Jak1	Janus kinase 1 (Jak1), mRNA [NM_146145]	-1.25	1.20	-1.05	-1.21	*1.97	**2.27	**2.08	*1.75
Arhgap22	Rho GTPase activating protein 22 (Arhgap22), mRNA [NM_153800]	-1.13	*1.35	-1.13	1.11	1.27	*1.65	**1.94	**2.27
Svil	supervillin (Svil), transcript variant 1, mRNA [NM_153153]	-1.26	1.37	1.03	-1.29	2.29	1.83	2.72	1.70
Irf1	interferon regulatory factor 1 (Irf1), mRNA [NM_008390]	*-1.34	**1.87	**2.12	-1.18	**2.26	*1.69	**1.86	**1.81
Add3	gamma-2 adducin (Add3) mRNA, alternatively spliced, complete cds. [AF100425]	-1.09	-1.07	1.68	*1.08	*-2.13	-1.37	*-2.23	**-2.26
Dusp10	dual specificity phosphatase 10	1.22	-1.11	*-2.13	*2.01	1.43	**2.18	*1.73	*1.84

	(Dusp10), mRNA [NM_022019]								
Cckbr	cholecystokinin B receptor (Cckbr), mRNA [NM_007627]	**-1.05	-1.06	-1.07	-1.05	**2.18	*1.50	*1.44	1.11
Psmb10	proteasome (prosome, macropain) subunit, beta type 10 (Psmb10), mRNA [NM_013640]	1.14	**1.93	1.10	1.11	**1.80	1.45	**2.16	**1.76
Ung2	0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:E130318K11 product:similar to URACIL-DNA GLYCOSYLASE 2 [Homo sapiens], full insert sequence [AK053890]	-1.12	**1.93	1.13	1.083	1.39	1.10	**2.16	**1.90
Dctn6	dynactin 6 (Dctn6), mRNA [NM_011722]	-1.09	*1.62	1.09	-1.07	**1.95	**2.02	**2.15	*1.69
Angptl6	angiopoietin-like 6 (Angptl6), mRNA [NM_145154]	1.06	1.37	-1.11	1.20	1.28	1.38	**1.80	**2.08
Rps6ka2	ribosomal protein S6 kinase, polypeptide 2 (Rps6ka2), mRNA [NM_011299]	-1.09	1.05	-1.29	-1.26	**1.97	1.14	**1.81	*1.56
2310016C08 Rik	RIKEN cDNA 2310016C08 gene (2310016C08Rik), mRNA	-1.15	1.17	-1.16	*1.69	*1.54	*1.69	**1.97	**1.91

	[NM_023516]								
Plod2	procollagen lysine, 2-oxoglutarate 5-dioxygenase 2 (Plod2), mRNA [NM_011961]	-1.19	*-1.40	1.22	-1.03	*-1.78	-1.13	**-1.96	**-1.74
H6pd	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase) (H6pd), mRNA [NM_173371]	1.26	**2.02	-1.19	-1.01	**1.95	1.21	**1.94	**1.69
Bok	Bcl-2-related ovarian killer protein (Bok), mRNA [NM_016778]	-1.05	-1.03	-1.16	1.10	*-1.40	-1.29	**-1.94	**-1.82
Btg2	B-cell translocation gene 2, anti-proliferative (Btg2), mRNA [NM_007570]	-1.20	1.12	-1.35	*1.72	**1.91	*1.62	*1.51	1.31
Bcor	Bcl6 interacting corepressor (Bcor), transcript variant a, mRNA [NM_029510]	1.11	1.15	-1.15	1.13	1.20	-1.05	**1.89	*1.51
Tec	cytoplasmic tyrosine kinase, Dscr28C related (Drosophila) (Tec), mRNA [NM_013689]	-1.01	-1.26	1.18	-1.06	*-1.35	-1.25	**-1.86	*-1.45
Hmgcr	3-hydroxy-3-methylglutaryl-Coenzyme A	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85

	reductase (Hmgcr), mRNA [NM_008255]								
Cbr3	carbonyl reductase 3 (Cbr3), mRNA [NM_173047]	1.20	*1.53	-1.23	1.09	**1.83	1.20	*1.58	**1.76
Fbxw4	F-box and WD-40 domain protein 4 (Fbxw4), mRNA [NM_013907]	1.00	*1.49	-1.05	1.07	*1.64	1.19	**1.75	**1.77
Rgnef	Rho-guanine nucleotide exchange factor (Rgnef), mRNA [NM_012026]	-1.04	*1.40	1.23	1.17	*1.53	*1.48	**1.70	**1.68
Nfkbia	nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha (Nfkbia), mRNA [NM_010907]	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Impdh1	inosine 5'-phosphate dehydrogenase 1 (Impdh1), mRNA [NM_011829]	1.13	1.06	-1.20	1.09	1.24	1.09	**1.56	**1.51
A_51_P144348	Unknown	-1.03	*1.74	1.14	1.25	*1.60	**1.71	**2.12	**2.20
2210016H18 Rik	adult male stomach cDNA, RIKEN full-length enriched library, clone:2210016H18 product:unclassifiable, full insert sequence [AK008737]	-1.03	*-1.42	*1.38	1.03	*-1.61	-1.29	*-1.92	**-2.04

Cited2	msg-related gene 1 (mrg1) mRNA, complete cds. [U86445]	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62
Svil	4 days neonate male adipose cDNA, RIKEN full-length enriched library, clone:B430302E16 product:SUPERVILLIN homolog [Homo sapiens], full insert sequence. [AK021019]	-1.26	1.37	1.03	-1.29	1.29	*1.83	**2.72	*1.70
Atxn2l	ataxin 2-like (Atxn2l), mRNA [NM_183020]	1.30	1.25	1.05	1.27	**-2.64	*-1.85	*-1.53	*-1.76
Lincr	lung-inducible neuralized-related C3HC4 RING domain protein (Lincr), mRNA [NM_153408]	1.10	1.26	-1.43	-1.01	1.57	1.11	**2.51	*1.74
Mfsd2	major facilitator superfamily domain containing 2 (Mfsd2), mRNA [NM_029662]	1.12	-1.31	1.08	1.00	**-2.54	*-2.65	-1.29	-1.41
Cdk6	adult male thymus cDNA, RIKEN full-length enriched library, clone:5830411I20 product:unclassifiable, full insert sequence [AK030810]	-1.10	*-1.56	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98

Phf17	PHD finger protein 17 (Phf17), mRNA [NM_172303]	1.25	*-1.69	1.05	-1.60	-1.19	1.10	-1.20	*-1.36
2900019G14 Rik	RIKEN cDNA 2900019G14 gene, mRNA (cDNA clone IMAGE:5706647). [BC078451]	-1.02	*2.19	1.32	-1.30	*2.61	**2.80	1.43	1.41
Id3	inhibitor of DNA binding 3 (Id3), mRNA [NM_008321]	*-1.51	-1.08	*1.74	1.15	*-1.99	**-2.51	-1.21	-1.48
Ext1	exostoses (multiple) 1 (Ext1), mRNA [NM_010162]	-1.09	*-1.64	-1.09	-1.07	*-1.67	*-1.47	*-1.50	**-1.91
Ngfb	nerve growth factor, beta (Ngfb), mRNA [NM_013609]	*1.97	1.43	*-1.97	1.34	*2.21	*1.46	1.41	*1.67

* significant $p<0.05$

** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml

D2- Concentration of coal tar used for exposure = 4 μg/ml

Tx- Hours of recovery time

(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation

x.xx down-regulation

5.4 Appendix D

Complete Ontology results

Table 5.3 Gene ontology categories created from genes that meet one of 3 criteria: significantly changed genes after FDR adjustment ($p<0.05$), significantly changed genes with unadjusted $p<0.001$, and/or genes showing a greater than 2-fold change in expression.

Gene Ontology: Functional annotation	Number of genes in category	percent of genes in pathway that are differentially expressed	Percent of genes in pathway that are differentially expressed – restricted those on Agilent chip	P-value	FDR
Sterol biosynthesis	*7*	*0.3*	*29.2*	*3.10E-06*	*0*

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Pmvk	phosphomevalonate kinase	-1.20	-1.01	-1.01	1.071	* -1.41	* -1.53	-1.11	** -2.03
Sc5d,	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.32	**-1.85
idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.98	*-2.16	*-1.65	**-3.86
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96

Nsdhl	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
Hmgcs1	3-hydroxy-3-methylglutaryl-coenzyme a synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.07
cholesterol biosynthetic process	*10*	*0.4*	*43.5*	*1.40E-08*	*0*				
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Pmvk	phosphomevalonate kinase	-1.20	-1.01	-1.01	1.07	*-1.41	*-1.53	-1.11	**-2.04
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.32	**-1.85
Idi1,	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.98	*-2.16	*-1.65	**-3.86
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Nsdhl	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
Hmgsc1	3-hydroxy-3-methylglutaryl-coenzyme a synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	*-1.83	**-3.07
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	* -1.43	-1.30	**-2.12
Dhcr7	7-dehdrocholesterol reductase	1.09	-1.08	-1.38	-1.43	-1.25	*-1.13	-2.13	*-2.00
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Dhcr24	24-dehydrocholesterol reductase	-1.07	1.06	-1.02	-1.08	-1.23	-1.37	-1.08	**-1.75
sterol biosynthetic process	*11*	*0.4*	*36.6*	*6.90E-09*	*0*				

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Pmvk	phosphomevalonate kinase	-1.20	-1.01	1.00	1.07	-1.41	* -1.54	* -1.11	** -2.04
Dhcr24	24-dehydrocholesterol reductase	-1.07	1.06	-1.02	-1.08	-1.72	*-1.40	-1.30	**- 3.63
Dhcr7	7-dehdrocholesterol reductase	1.09	-1.08	-1.38	-1.43	-1.25	*-1.13	-2.13	*-2.00
Sc5d	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**- 2.20
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	-1.85
Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.98	*-2.16	*-1.65	**- 3.86
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	*-1.43	-1.30	**- 2.12
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	** -4.47
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Nsdhl	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
Hmgcs1	3-hydroxy-3-methylglutaryl-coenzyme a synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	*-1.83	**- 3.08
biosynthesis of steroids	*12*	*0.5*	*30.8*	*4.60E-11*	*0*				

Gene symbol	Gene name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8

Pmvk	phosphomevalonate kinase	-1.20	-1.01	1.00	1.07	*-1.41	*-1.54	-1.11	**-2.04
Sqle	squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	*-1.43	**-2.65
Dhcr24	24-dehydrocholesterol reductase	1.07	1.13	1.03	1.32	-1.72	-1.40	-1.30	*-3.63
Dhcr7	7-dehydrocholesterol reductase	1.09	-1.08	-1.38	-1.43	-1.25	*-1.13	-2.13	*-2.00
Nqo1	nad(p)h dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36
Sc5d,	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20
Hmgcs1	3-hydroxy-3-methylglutary-coenzyme A synthase	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.08
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85
Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	*-1.43	-1.30	**-2.12
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Lss	lanosterol synthase	-1.21	-1.18	-1.03	1.17	*-1.80	*-2.05	-1.39	**-3.03
Nsdhl,	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
steroid metabolic process	*8*	*0.7*	*5.2*	*5.90E-07*	*0*				
Gene symbol	**Gene name**	**Fold change**							

		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Sc5d	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85
idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Pmvk	phosphomevalonate kinase	-1.20	-1.01	1.00	1.07	*-1.41	*-1.54	-1.11	**-2.04
Stard4	riken cDNA 4632419c16 gene	1.19	-1.30	1.02	1.02	*-2.08	*-1.82	*-1.64	*-1.90
Dhcr24	24-dehydrocholesterol reductase	1.07	1.13	1.03	1.32	-1.72	-1.40	-1.30	**-3.63
Dhcr7	7-dehdrocholesterol reductase	1.09	-1.08	-1.38	-1.43	-1.25	*-1.13	-2.13	*-2.00
regulation of cell differentiation	*13*	*0.7*	*7.6*	*4.20E-06*	*0*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Nfkbia	nuclear factor of kappa light chain gene enhancer in b-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Tgfb2	follistatin	1.08	*-1.58	1.13	-1.13	-1.44	* -2.15	* -1.83	* -1.91

Foxg1	transforming growth factor, beta 2	1.17	*-1.60	*-1.64	1.02	-1.30	1.06	-1.27	1.04
Runx1	runt related transcription factor 1	1.71	-1.10	-1.61	1.26	*1.66	*2.07	1.35	1.04
Mapk1	mitogen activated protein kinase 1	-1.01	-1.29	1.18	-1.27	*-1.33	1.08	-1.20	-1.59
Bmp4	bone morphogenetic protein 4	1.27	*-1.60	-1.47	-1.30	**-2.24	*-1.96	-1.16	*-1.56
Foxg1	forkhead box g1	1.17	*-1.60	-1.64	1.02	-1.30	1.06	-1.27	1.04
hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	*-2.49	-1.15	*-1.82	-1.60
Il6	interleukin 6	1.69	-1.15	-1.05	**2.86	-1.35	-1.31	*-1.85	*-2.03
Nrg1	neuregulin 1	1.21	1.27	-1.16	1.04	*1.60	1.30	*1.39	1.02
wwtr1	ww domain containing transcription regulator 1	-1.49	*-2.15	-1.09	-1.13	-1.08	1.29	-1.45	-1.57
Gli3	gli-kruppel family member gli3	-1.06	*-1.26	1.26	1.00	-1.20	* -1.10	1.01	-1.14
Xdh	xanthine dehydrogenase	1.04	1.09	-1.06	1.05	-1.05	1.15	**2.48	*2.19
angiogenesis	*17*	*0.7*	*13.5*	*7.60E-08*	*0*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Ctgf	connective tissue growth factor	1.02	-1.45	-2.11	1.01	1.11	1.01	-1.08	1.13
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	*1.80
Ubp1	upstream binding protein 1	-1.21	1.01	*2.20	1.42	-1.31	-1.03	1.12	1.14
Edn1	endothelin 1	1.38	**-1.47	-1.41	-1.13	*-1.73	-1.14	*-2.05	-1.42
Dicer1	dicer1, dcr-1 homolog (drosophila)	-1.19	-1.40	-1.07	-1.18	-1.04	1.28	-1.23	* -2.03

Runx1	runt related transcription factor 1	1.71	-1.10	-1.61	1.26	*1.66	*2.07	1.35	1.04
Angptl4	angiopoietin-like 4	1.87	-1.01	-2.12	-1.17	*1.63	1.21	1.25	*1.51
Rtn4	reticulon 4	-1.16	-1.21	1.10	-1.04	*-2.18	-1.51	-1.22	-1.35
Bmp4	bone morphogenetic protein 4	1.27	-1.60	-1.47	-1.30	**-2.24	*-1.96	-1.16	*-1.56
Fgf18	fibroblast growth factor 18	1.54	-1.10	-1.12	-1.11	-2.24	*-1.96	-1.16	**-1.56
Hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	*-2.49	-1.15	*-1.82	-1.60
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Hbegf	heparin-binding egf-like growth factor	1.91	-1.50	-2.68	-1.01	*-1.73	-1.27	*-1.71	*-1.86
Angptl6	angiopoietin-like 6	1.06	1.37	-1.11	1.20	1.28	1.38	**-1.02	**1.18
Adamts1	a disintegrin-like and metallopeptidse (reprolysin type) with thrombospondin type 1 motif, 1	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	*-1.88	**-2.44
Arhgap22	rho GTPase activating protein 22	-1.13	*1.35	-1.13	1.11	1.27	*1.65	**1.94	**2.27
lipid biosynthetic process	*22*	*0.8*	*14.9*	*4.80E-06*	*0*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Sc5d	sterol-c5-desaturase (fungal erg3, delta-5-desaturase) homolog (s. cerevisae)	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85

Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86
Ptgs1	prostaglandin-endoperoxide synthase 1	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	*1.99	*2.67
Hsd17b7	hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	*-1.96	*-1.72	*-1.66	**-4.47
Mvd	mevalonate (diphospho) decarboxylase	1.01	1.04	-1.55	-1.30	-1.16	*-2.23	1.18	*-1.96
Pmvk	phosphomevalonate kinase	-1.20	-1.01	1.00	1.07	*-1.41	*-1.54	-1.11	**-2.04
Stard4	riken cDNA 4632419c16 gene	1.19	-1.30	1.02	1.02	*-2.08	*-1.82	*-1.64	*-1.90
Pcx	pyruvate carboxylase	1.22	1.21	-1.15	1.11	*1.42	1.01	*1.60	*1.30
Dhcr24	24-dehydrocholesterol reductase	1.07	1.12	1.03	1.32	-1.72	-1.40	-1.30	**-3.63
Dctn6	dynactin 6	-1.09	*1.62	1.09	-1.07	**1.95	**2.02	**2.15	*1.70
Scd2	stearoyl-coenzyme a desaturase 2	-1.16	-1.24	1.08	-1.39	*-1.65	**-2.06	1.03	**-2.20
Ptgs2	prostaglandin-endoperoxide synthase 2	*2.26	-1.53	*-3.29	1.59	*-1.65	*1.52	*-1.15	**-1.39
Pcyt1a	phosphate cytidylyltransferase 1, choline, alpha isoform	1.10	*-1.32	1.24	1.18	-1.20	1.09	-1.24	*-1.57
Cyp51	cytochrome p450, family 51	-1.45	-1.45	1.08	1.10	-1.33	-1.43	-1.30	**-2.12
Lss	lanosterol synthase	-1.21	-1.18	-1.03	1.17	*-1.80	*-2.05	-1.39	**-3.03
Nsdhl	nad(p) dependent steroid dehydrogenase-like	1.17	-1.14	-1.09	1.01	*-1.07	-1.01	1.03	*1.05
Hmgcs1	3-hydroxy-3-methylglutaryl-coenzyme a synthase 1	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.08
Elovl6	elovl family member 6, elongation of long chain fatty acids (yeast)	-1.27	-1.25	1.56	-1.45	-1.31	*-1.54	*-1.51	**-2.90

tissue development	*22*	*1*	*16.4*	*9.20E-07*	*0*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ext1	exostoses (multiple) 1	-1.09	*-1.64	-1.09	-1.07	*-1.67	*-1.47	*-1.50	**-1.91
Mgp	matrix gla protein	1.35	-1.24	-1.26	-1.12	*-1.54	*-1.88	*-2.18	*-2.04
Ctgf	connective tissue growth factor	1.02	*-1.45	-2.11	1.014	1.11	1.01	-1.08	1.13
Fst	follistatin	1.62	-1.29	-2.13	-1.14	-1.27	1.02	-1.29	**-2.30
Tgfb2	transforming growth factor, beta 2	1.08	*-1.58	1.13	-1.13	-1.44	*-2.15	*-1.83	*-1.91
Gja1	gap junction membrane channel protein alpha 1	-1.01	-1.34	1.02	-1.01	-1.44	-1.06	*-1.78	**-2.42
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.68	**-2.31
Bmp2k	riken cDNA 4933417m22 gene	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*1.69	-1.43
Edn1	endothelin 1	1.38	*-1.47	-1.41	-1.13	*-1.73	-1.14	*-2.05	-1.42
Adrb2	adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	*2.02	**2.12
Traf6	tnf receptor-associated factor 6	2.32	1.19	-1.54	1.29	1.00	-1.26	-1.18	1.12
Bmp4	bone morphogenetic protein 4	1.27	*-1.60	-1.47	-1.30	**-2.24	*-1.96	-1.16	*-1.56
Snai1	snail homolog 1 (drosophila)	1.15	-1.32	**-2.46	-1.19	-1.35	-1.25	1.44	1.12
Dhcr24	24-dehydrocholesterol reductase	1.07	1.13	1.03	1.32	-1.72	-1.40	-1.30	**-3.63

blood vessel development/vascular development	*23*	*1*	*9.8*	*1.10E-09*	*0*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	1.40	*-1.77	**-2.79	1.35	*-1.52	-1.18	-1.32	*-1.69
Ctgf	connective tissue growth factor	1.02	*-1.45	-2.11	1.01	1.11	1.01	-1.08	1.13
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.80
Tgfb2	transforming growth factor, beta 2	1.08	*-1.58	1.13	-1.13	-1.44	*-2.15	*-1.83	*-1.91
Gja1	gap junction membrane channel protein alpha 1	-1.01	-1.34	1.02	-1.01	-1.44	-1.06	*-1.78	**-2.43
Ubp1	upstream binding protein 1	-1.21	1.01	2.20	1.42	*-1.31	-1.03	1.12	1.14
Dicer1	dicer1, dcr-1 homolog (drosophila)	1.05	-1.20	1.12	1.11	-1.39	-1.08	-1.02	-1.22
Edn1	endothelin 1	1.38	*-1.47	-1.41	-1.13	*-1.73	-1.14	*-2.05	-1.42
Runx1	runt related transcription factor 1	1.71	-1.10	-1.61	1.26	*1.66	*2.07	1.35	1.04
Angptl4	angiopoietin-like 4	1.87	-1.01	-2.12	-1.17	*-1.01	-1.19	-1.13	*-1.02
Rtn4	reticulon 4	-1.16	-1.21	1.10	-1.04	*-2.18	-1.51	-1.22	-1.35
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Cited2	cbp/p300-interacting transactivator, with glu/asp-rich carboxy-terminal domain, 2	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62

Bmp4	bone morphogenetic protein 4	1.27	*-1.60	-1.47	-1.30	**-2.24	* -1.96	-1.16	*-1.57
Aw125753	riken cDNA 4731402f03 gene	1.83	-1.36	-1.11	-1.03	* -1.74	* -1.86	-1.72	** -2.66
Hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	*-2.49	-1.15	*-1.82	-1.60

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

Table 5.4 Gene ontology categories created from the list of genes that showed significant expression changes in at least one condition, after FDR adjustment. (Note: Not all conditions are shown)

Gene Ontology: Functional annotation	Number of genes in category	percent of genes in pathway that are differentially expressed	Percent of differentially expressed genes in pathway vs total genes in pathway from the chip	P-value	FDR
Oxidoreductase activity	*16*	*2.8*	*2.9*	*5.110E-05*	*0*

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Dhrs7	dehydrogenase/reductase (sdr family) member 7	1.05	**1.74	1.12	1.02	**2.12	*2.22	**2.09	**2.64
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	*1.80
Nqo1	nad(p)h dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36
H6pd	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)	1.26	**2.02	-1.19	-1.01	**1.95	1.21	**1.94	**1.69
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.32	-1.85

	reductase								
Adh7	alcohol dehydrogenase 7 (class iv), mu or sigma polypeptide	-1.13	*1.38	-1.01	-1.10	*1.97	*2.08	2.99	2.36
Ptgs1	prostaglandin-endoperoxide synthase 1	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	1.99	2.67
Cpox	coproporphyrinogen oxidase	*-1.43	**2.46	1.15	-1.35	**3.29	**3.65	**3.82	**3.63
Aldh3a1	aldehyde dehydrogenase family 3, subfamily a1	1.00	*1.43	-1.28	1.298	2.03	1.41	3.08	2.24
Plod2	procollagen lysine, 2-oxoglutarate 5-dioxygenase 2	-1.19	*-1.40	1.22	-1.03	*-1.78	-1.13	-1.96	-1.74
Sqle	squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	-1.43	-2.65
Cyp1a1	cytochrome p450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	cytochrome p450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Cbr3	carbonyl reductase 3	1.20	*1.53	-1.23	1.09	**1.83	1.20	1.56	1.76
Impdh1	inosine 5'-phosphate dehydrogenase 1	1.13	1.06	-1.20	1.09	1.24	1.09	**1.56	**1.51
Dhrs7	dehydrogenase/reductase (sdr family) member 7	1.05	**1.74	1.12	1.02	**2.12	*2.23	**2.09	**2.64

Oxidoreductase activity, acting on the CH-OH, group of donors, NAD or NADP as an acceptor	*6*	*1*	*23.1*	*3.0E-04*	*0.1*				
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
H6pd	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)	1.26	**2.02	-1.19	-1.01	**1.95	1.21	**1.94	**1.69
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	-1.85
Adh7	alcohol dehydrogenase 7 (class iv), mu or sigma polypeptide	-1.13	*1.38	-1.01	-1.10	*1.97	*2.08	2.99	2.36
Aldh3a1	aldehyde dehydrogenase family 3, subfamily a1	1.00	*1.43	-1.28	1.298	2.03	1.41	3.08	2.24
Cbr3	carbonyl reductase 3	1.20	*1.53	-1.23	1.09	**1.83	1.20	1.59	1.76
Impdh1	inosine 5'-phosphate dehydrogenase 1	1.13	1.06	-1.20	1.09	1.24	1.09	**1.56	**1.51
Developmental	*32*	*5.6*	*13.2*	*5.90E-05*	*0.5*				

process, multicellular organismal development, cellular differentiation									
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Tnfrsf12a	tumor necrosis factor receptor superfamily, member 12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Ngfb	nerve growth factor, beta	*1.97	1.43	*-1.97	1.34	*2.21	*1.46	1.41	*1.67
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme a reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	-1.85
Rgnef	rho interacting protein 2	-1.04	*1.40	1.23	1.17	*1.53	*1.48	**1.70	1.68
Btg2	b-cell translocation gene 2, anti-proliferative	-1.20	1.12	-1.35	*1.72	**1.91	*1.62	1.51	1.31
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.68	-2.31
Adrb2	adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	2.02	2.12
Prrx2	paired related homeobox 2	-1.03	*1.35	*-1.50	-1.02	*1.70	*1.84	2.89	2.44
Nr4a1	nuclear receptor	*2.07	*-1.45	**-5.71	**2.25	1.00	1.05	1.01	1.28

	subfamily 4, group a, member 1								
Fcgr3	fc receptor, IgG, low affinity iii	-1.08	**2.19	1.08	1.03	1.45	*1.76	*1.86	*1.65
Phf17	phd finger protein 17	1.25	*-1.69	1.05	-1.60	-1.19	1.10	-1.20	*-1.36
Fbxw4	f-box and wd-40 domain protein 4	1.00	1.49	-1.05	1.07	1.64	1.19	1.75	1.77
Adamts1	a disintegrin-like and metallopeptidse (reprolysin type) with thrombospondin type 1 motif, 1	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	-1.88	-2.44
Angptl6	angiopoietin-like 6	1.06	1.37	-1.11	1.20	1.28	1.38	**1.80	2.08
Myd116	myeloid differentiation primary response gene 116	-1.30	-1.23	-1.31	**2.26	1.06	*1.67	1.29	1.38
Nfkbia	nuclear factor of kappa light chain gene enhancer in b-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Ext1	exostoses (multiple) 1	-1.09	*-1.64	-1.09	-1.07	*-1.67	*-1.47	*-1.50	**-1.91
Shroom1	riken cDNA 1300007I22 gene	1.19	*1.54	*-1.66	1.32	**2.68	*1.75	1.76	1.54
Wisp2	wnt1 inducible signaling pathway protein 2	1.09	*1.88	-1.46	-1.01	**2.70	1.45	3.22	2.27
Bok	bcl-2-related ovarian killer protein	-1.05	-1.03	-1.16	1.10	*-1.40	-1.29	-1.94	-1.82
Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	* 1.80
Gpc1	glypican 1	-1.03	**2.50	*-1.45	1.12	**2.45	**2.65	**4.17	**4.48
Ptgs1	prostaglandin-endoperoxide synthase 1	1.17	*1.58	*-1.45	-1.19	*2.14	*1.64	1.99	2.67
Ntrk2	neurotrophic tyrosine kinase, receptor, type	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49

	2								
Id3	inhibitor of DNA binding 3	*-1.51	-1.08	*1.74	1.15	*-1.99	**-2.51	-1.21	-1.48
Cited2	cbp/p300-interacting transactivator, with glu/asp-rich carboxy-terminal domain, 2	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62
Spry1	sprouty homolog 1 (drosophila)	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	2.30	1.81
Rps6ka2	ribosomal protein s6 kinase, related sequence 1	-1.09	1.05	-1.29	-1.26	**1.97	1.14	1.81	1.56
Dfna5h	deafness, autosomal dominant 5 homolog (human)	1.53	1.54	-1.31	1.03	**2.34	1.42	**3.38	3.32
Ifrd1	interferon-related developmental regulator 1	-1.15	-1.01	1.27	*1.85	1.08	**2.89	-1.01	1.25
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Arhgap22	rho GTPase activating protein 22	-1.13	*1.35	-1.13	1.11	1.27	*1.65	1.94	2.27
Blood vessel development, vascular development, angiogenesis	*8*	*1.4*	*3.4*	*1.7E-04*	*0.3*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
	tumor necrosis factor receptor superfamily, member 12a	*1.40	*-1.77	**-2.79	1.35	-1.51	-1.18	-1.32	*-1.69

Nos3	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	* 1.80
Id1	inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Adamts1	a disintegrin-like and metallopeptidse (reprolysin type) with thrombospondin type 1 motif, 1	**2.36	*-1.53	**-2.95	1.21	*-1.74	*-1.50	-1.88	-2.44
Angptl6	angiopoietin-like 6	1.06	1.37	-1.11	1.20	1.28	1.38	1.79	2.08
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Arhgap22	rho GTPase activating protein 22	-1.13	*1.35	-1.13	1.11	1.27	*1.65	1.94	2.27
Cited2	cbp/p300-interacting transactivator, with glu/asp-rich carboxy-terminal domain, 2	**2.00	-1.17	*-1.71	1.21	*-2.03	-1.44	*-1.48	**-1.62
metalloprotein	*5*	*0.9*	*9.25*	*2.30E-04*	*0.4*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
	nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	** 2.29	* 1.80
Adh7	alcohol dehydrogenase 7 (class iv), mu or sigma polypeptide	-1.13	*1.38	-1.01	-1.10	*1.97	*2.08	2.99	2.36
Cyp1a1	cytochrome p450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cpox	coproporphyrinogen	*-1.43	**2.46	1.15	-1.35	**3.29	**3.65	**3.82	**3.63

	oxidase								
Cyp1b1	cytochrome p450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Biosynthesis of steroids	*5*	*0.9*	*12.8*	*2.30E-04*	*0.4*				
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Sqle	squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	-1.42644	-2.65
Nqo1	nad(p)h dehydrogenase, quinone 1	-1.11	1.92	*-1.43	-1.44	**2.68	*2.05	**3.35	**3.36
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31524	-1.85
Hmgcs1	3-hydroxy-3-methylglutary-coenzyme A synthase	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	-1.83	**-3.08
Idi1	isopentenyl-diphosphate delta isomerase	-1.56	*-1.57	1.58	-1.37	*-1.978	*-2.16	*-1.65	**-3.86

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)

x.xx up-regulation
x.xx down-regulation

Table 5.5 Results of KEGG Pathway analyses employing genes that meet one of the following criteria (1) significantly changed genes with FDR adjusted p<0.05, (2) significantly changed genes with unadjusted p<0.001, or (3) genes showing a greater than 2-fold change.

KEGG Pathway	**Number of differentially expressed genes in category**	**Number of genes in pathway**	**Pathway identifier**						
Cancer	*19*	*345*	*mmu05200*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Bmp4	bone morphogenetic protein 4	1.27	-1.6	-1.47	-1.3	-2.24	-1.96	-1.16	-1.56
Cdk6	cyclin-dependent kinase 6	-1.1	-1.02	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Fgf18	fibroblast growth factor 18	1.54	-1.1	-1.12	-1.11	-1.46	-2.21	-1.4	-2.7
Fgf7	fibroblast growth factor 7	1.82	1	-2.08	1.02	-1.63	-1.1	-1.2	1
Fzd4	frizzled homolog 4	1.14	-1.13	1.06	1.06	-1.04	-1.44	-2.33	-2.41
Gli3	GLI-Kruppel family member	-1.26	-2.03	1.31	1.35	-1.2	-1.1	1.01	-1.14
Hif1a	hypoxia inducible factor 1, alpha subunit	1.06	-1.46	1.38	-1.42	-2.49	-1.15	-1.82	-1.6
Il6	interleukin 6	1.69	-1.15	-1.05	2.86	-1.35	-1.31	-1.85	-2.03
Jak1	Janus kinase 1	-1.25	1.2	-1.05	-1.21	*1.97	2.27	2.08	1.75
Kitl	kit ligand	-1.09	-1.18	1.24	1.08	1.15	2	1.13	1.1
Mapk1	mitogen-activated protein kinase 1	-1.4	-1.4	1.12	1	-1.33	1.08	-1.2	-1.59
Myc	myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60

	inhibitor, alpha								
Nos3	nitric oxide synthase 3	1.12	1.22	-1.13	1.14	1.29	*1.85	**	*
Ptgs2	prostaglandin-endoperoxide synthase 2	2.26	-1.53	-3.29	1.59	-1.65	1.52	2.29	1.8
Ralbp1	ralA binding protein 1	-1.27	-1.62	-1.07	-2.02	1.29	1.2	-1.15	-1.39
Runx1	runt related transcription factor 1	1.71	-1.1	-1.61	1.26	1.66	2.07	-1.14	1.01
Tgfb2	transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	1.35	1.04
Traf6	TNF receptor-associated factor 6	1.01	-1.03	1.09	1.08	-1.04	-1.12	-1.83	-1.91
MapK signalling pathway	*14*	*286*	*Mmu04010*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ddit3	DNA-damage inducible transcript 3	-1.63	-1.07	1.5	1.58	1.29	2.15	1.22	1.38
Dusp10	dual specificity phosphatase 10	1.22	-1.11	*-2.13	*2.01	1.43	**2.18	1.73	1.84
Dusp4	dual specificity phosphatase 4	2.39	1.04	-2.91	1.24	1.18	1.23	1.15	-1.13
Dusp6	dual specificity phosphatase 6	1.51	-1.41	-3.42	-1.49	-1.1	-1.11	-1.19	-1.2
Fgf18	fibroblast growth factor 18	1.54	-1.1	-1.12	-1.11	-1.46	-2.21	-1.4	-2.7
Fgf7	fibroblast growth factor 7	1.82	1	-2.08	1.02	-1.63	-1.1	-1.2	1

Gadd45a	growth arrest and DNA-damage-inducible 45 alpha	-1.83	-1.1	1.65	1.19	1.32	2.11	1.3	1.61
Mapk1	Mitogen-activated protein kinase receptor	-1.17	-1.4	1.12	1	-1.33	1.08	-1.2	-1.59
Myc	Myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	*2.07	*-1.45	**-5.71	**2.25	1	1.05	1.01	1.28
Ntrk2	Neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Rps6ka2	Ribosomal protein S6 kinase, polypeptide 2	-1.09	1.05	-1.29	-1.26	**1.97	1.14	1.81	1.56
Tgfb2	Transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	-1.83	-1.91
Traf6	TNF receptor-associated factor 6	1.01	-1.03	1.09	1.08	-1.04	-1.12	1.09	1.26
Cytokine-cytokine receptor interaction	*12*	*272*	*Mmu04060*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cc12	chemokine (C-C motif) ligand 2	2.39	-1.81	-2.25	-1.04	-2.49	-2.32	-1.49	-2.46
Cc17	chemokine (C-C motif) ligand 7	3.22	-2.18	-3.07	-1.11	-2.07	-2.81	-3	-3.36
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	-1.68	-2.31

Csf2rb2	colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)	1.05	1.21	-1.11	1.1	-1.35	-2.05	1.12	1.25
Cxcl16	chemokine (C-X-C motif) ligand 16	2.04	2.04	-1.21	-1.1	1.24	1.31	1.15	1.05
Il6	Interleukin 6	1.69	-1.15	-1.05	2.86	-1.35	-1.31	-1.85	-2.03
Inhbb	Inhibin beta-B	-1.07	-1.23	-1.3	-1.18	1.51	1.1	1.94	2.5
Kitl	Kit ligand	-1.09	-1.18	1.24	1.08	1.15	2	1.13	1.1
Tgfb2	Transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	-1.83	-1.91
Tnfrsf12a	Tumor necrosis factor receptor superfamily, member12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Tnfrsf19	Tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.01	**1.26	*-1.20	1.1
Tslp	Thymic stromal lymphopoietin	1.44	1.41	1	1.03	1.2	1.67	1.37	1.34
TGF-beta signalling	*11*	*92*	*Mmu04350*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Id1	Inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Id3	Inhibitor of DNA binding 3	*-1.51	-1.08	*1.74	1.15	*-1.99	**-2.51	-1.21	-1.48
Inhbb	Inhibin beta-B	-1.07	-1.24	-1.3	-1.18	1.51	1.1	1.94	2.5
Mapk1	Mitogen-activated protein kinase 1	-1.17	-1.4	1.12	1	-2.02	-1.25	-1.46	-1.14

Myc	Myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Rock2	Rho-associated coiled-coil containing protein kinase 2	-1.08	-1.37	1.29	1.19	-1.48	-1.15	-1.81	-1.5
Smad7	MAD homolog 7	1.01	-1.74	-1.59	-1.24	-1.85	-1.24	-1.75	-2.38
Tgfb2	Transforming growth factor, beta 2	1.08	-1.58	1.13	-1.13	-1.44	-2.15	-1.83	-1.91
Chemokine signalling	*10*	*203*	*mmu04062*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adcy1	Adenyate cyclase 1	-1.03	1.16	-1.41	-1.12	1.23	2.1	1.33	-1.1
Ccl2	Chemokine (C-C motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	-1.49	-2.46
Ccl20	Chemokine (C-C motif) ligand 20	1.2	1.14	-1.44	1.48	1.66	2.32	1.79	1.22
Ccl7	Chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Cxcl16	Chemokine (C-X-C motif) ligand 16	2.04	2.04	-1.21	-1.1	1.24	1.31	1.15	1.05
Gnb4	guanine nucleotide binding protein (G protein), beta 4	1.29	*2.03	-1.11	-1.18	*1.88	**2.85	2.58	2.25
Mapk1	Mitogen-activated protein kinase 1	-1.17	-1.4	1.12	1	-1.33	1.08	-1.2	-1.59
Nfkbia	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60

Rock2	Rho-associated coiled-coil containing protein kinase 2	-1.08	-1.37	1.29	1.19	-1.48	-1.15	-1.81	-1.5
Shc4	Src homology 2 domain containing family, member 4	1	1.05	1.14	1.09	-1.2	-1.18	1.25	1.14
Neuroactive ligand-receptor interaction	*7*	*332*	*Mmu04080*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adrb2	Adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	2.02	2.12
Cckbr	Cholecystokinin B receptor	**-1.05	-1.06	-1.07	-1.05	**2.18	1.5	1.44	1.11
Cysltr1	Cysteinyl leukotriene receptor 1	-1.01	-1.21	1.46	-1.25	1.01	1.01	-1.55	-2.22
Gpr83	G protein-coupled receptor 83	1.07	-1.02	1.32	1.2	-1.74	-2.13	-1.21	-1.07
Mc1r	Melanocortin 1 receptor	1.43	1.39	1.3	1.62	1.07	-1.18	1.17	1.02
P2rx2	Purinergic receptor P2X, ligand-gated ion channel, 2	1.7	1.57	1	1.28	1.12	-1.02	-1.04	-1.06
Prss2	Protease	-1.32	-1.44	-1.19	-1.3	2.13	1.12	1.15	-1.17
Neurotrophin signalling	*7*	*126*	*Mmu04722*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8

Irak2	interleukin-1 receptor-associated kinase 2	1.47	1.75	-1.55	1.25	1.9	1.55	1.96	1.79
Mapk1	mitogen-activated protein kinase 1	-1.17	-1.4	1.12	1	-2.02	-1.25	-1.46	-1.14
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Ntrk2	Neurotrophic tyrosine kinase	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Rps6ka2	Ribosomal protein s6 kinase, polypeptide 2	-1.09	1.05	-1.29	-1.26	**1.97	1.14	**1.81	1.56
Shc4	Src homology 2 domain containg family, member 4	1	1.05	1.14	1.09	-1.2	-1.18	1.25	1.14
Traf6	TNF receptor-associated factor 6	1.01	-1.03	1.09	1.08	-1.04	-1.12	1.09	1.26
Ubiquitin mediated proteolysis	*7*	*156*	*Mmu04120*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cul5	Cullin 5	-1.25	-1.49	1.08	-2.19	1.05	1.21	-1.22	1.03
Klhl13	Kelch-like 13	-1.11	1.13	2.28	1.02	1.12	1.07	-1.57	-2.1
Nedd41	Neural precursor cell expressed, developmentally down-regulated gene 4-like	1.04	-1.08	1.12	1.12	-2.16	-1.29	1.13	1.12
Socs3	Suppressor of cytokine signalling 3	2.06	1.3	-1.17	1.08	1.17	-1.03	1.29	1.18
Traf6	TNF receptor-associated factor 6	1.01	-1.03	1.09	1.08	-1.04	-1.12	1.09	1.26

Ube1l	Ubiquitin-activating enzyme e1-like	1.11	1.31	-1.22	1.1	1.73	1.01	2.09	1.77
Ube2s	Ubiquitin-conjugating enzyme E2S	1.14	1.17	-1.36	1.31	-1.91	-1.74	1.05	1.23
Jak-STAT signaling	*7*	*160*	*Mmu 4630*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Csf2rb2	colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)	1.05	1.21	-1.11	1.1	-1.35	-2.05	1.12	1.25
Il6	Interleukin 6	1.69	-1.15	-1.05	2.86	-1.35	-1.31	-1.85	-2.03
Jak1	Janus kinase 1	-1.25	1.2	-1.05	-1.21	*1.97	**2.27	2.08	1.75
Myc	Myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Socs3	Suppressor of cytokine signalling 3	2.06	1.3	-1.17	1.08	1.17	-1.03	1.29	1.18
Spry1	Sprouty homolog 1	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	2.3	1.81
Tslp	Thymic stromal lymphopoietin	1.44	1.41	1	1.03	1.2	1.67	1.37	1.34
Steroid biosynthesis	*7*	*17*	*Mmu 00100*						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cyp51	Cytochrome P450, family 51	-1.45	-1.45	1.08	1.1	-1.33	-1.43	-1.3	-2.12

Dhcr24	24-dehydrocholesterol reductase	-1.07	1.06	-1.02	-1.08	-1.23	-1.37	-1.08	-1.75
Hsd17b7	Hydroxysteroid (17-beta) dehydrogenase 7	-1.18	-2.49	-1.62	-1.06	-1.96	-1.72	-1.66	-4.47
Lss	Lanosterol synthase	-1.21	-1.18	-1.03	1.17	-1.8	-2.05	-1.39	-3.03
Nsdh1	NAD(P) dependent steroid dehydrogenase 7	1.17	-1.14	-1.09	1.01	-1.07	-1.01	1.03	1.05
Sc5d	Sterol-C5-desturase	-1.27	-1.29	1.12	-1.23	-1.21	*-1.43	*-1.76	**-2.20
Sqle	Squalene epoxidase	-1.02	*-1.58	-1.23	-1.18	*-1.83	*-2.11	-1.43	-2.65
Small cell lung cancer	6	94	Mmu05222						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cdk6	Cyclin-dependent kinase 6	-1.1	*-1.56	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Myc	Myelocytomatosis oncogene	1.79	1.04	-2.58	1.16	-1.08	1.04	1.07	1.08
Nfkbia	Nuclear factor of kappa light polypeptide geneenhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Nos3	Nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**	*
Ptgs2	Prostaglandin-endoperoxide synthase 2							2.29	1.8
Traf6	TNF receptor-associated factor 6	2.26	-1.53	-3.29	1.59	-1.65	1.52	-1.15	-1.39

Metabolism of xenobiotics by cytochrome P450	5	30	Mmu0980						
Gene symbol	**Gene name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adh7	Alcohol dehydrogenase 7 (class IV)	-1.13	*1.38	-1.01	-1.1	*1.97	*2.08	2.99	2.36
Aldh3a1	Aldehyde dehydrogenase family 3, subfamily A1	1	*1.43	-1.28	1.3	2.03	1.41	3.08	2.24
Cyp1a1	Cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	Cytochrome P450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Gstm6	Glutathione S-transferase, mu 6	1.76	1.39	1.26	1.38	-1.66	-1.23	-1.22	-1.7

* significant $p<0.05$
** significant FDR adjusted $p<0.05$

D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)
x.xx up-regulation
x.xx down-regulation

Table 5.6 Results of KEGG Pathway analyses created from the FDR adjusted p<0.05 gene list.

KEGG Pathway	**Number of genes in category**	**Pathway identifier**							
Cytokine-cytokine receptor interaction	*5*	*Mmu04060*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cc12	chemokine (C-C motif) ligand 2	2.39	-1.81	-2.25	-1.04	**-2.49	**-2.32	*-1.49	**-2.46
Cc17	chemokine (C-C motif) ligand 7	3.22	-2.18	-3.07	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Csf1	colony stimulating factor 1 (macrophage)	1.31	**-2.00	*-1.99	*-1.38	**-2.23	*-2.60	*-1.67	-2.31
Tnfrsf12a	Tumor necrosis factor receptor superfamily, member12a	*1.40	*-1.77	**-2.79	1.35	-1.52	-1.18	-1.32	*-1.69
Tnfrsf19	Tumor necrosis factor receptor superfamily, member 19	-1.21	**2.83	*1.80	1.06	1.01	**1.26	*-1.20	1.1
cancer	*4*	*Mmu05200*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8

Cdk6	cyclin-dependent kinase 6	-1.1	*-1.56	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Jak1	Janus kinase 1	-1.25	1.2	-1.05	-1.21	*1.97	**2.27	**2.08	*1.75
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Nos3	nitric oxide synthase 3	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
MapK signalling pathway	*4*	*Mmu04010*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Dusp10	dual specificity phosphatase 10	1.22	-1.11	*-2.13	*2.01	1.43	**2.18	*1.73	*1.84
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	*2.07	*-1.45	**-5.70	**2.25	1	1.05	1.01	1.28
	Neurotrophic tyrosine kinase, receptor, type 2	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Ntrk2									
Rps6ka2	Ribosomal protein S6 kinase, polypeptide 2	-1.09	1.05	-1.29	-1.26	**1.97	1.14	**1.81	*1.56
Chemokine signalling	*4*	*mmu04062*							
Gene symbol	**Gene Name**	**Fold change**							

		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ccl2	Chemokine (C-C motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	*-1.49	**-2.46
Ccl7	Chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Gnb4	guanine nucleotide binding protein (G protein), beta 4	1.29	*2.03	-1.11	-1.18	*1.88	**2.85	**2.58	**2.25
Metabolism of xenobiotics by cytochrome P450	*4*	*Mmu0980*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adh7	Alcohol dehydrogenase 7 (class IV)	-1.13	*1.38	-1.01	-1.1	*1.97	*2.08	**2.99	*2.36
Aldh3a1	Aldehyde dehydrogenase family 3, subfamily A1	1	*1.43	-1.28	1.298	2.03	1.41	**3.08	**2.24
Cyp1a1	Cytochrome P450, family 1, subfamily a, polypeptide 1	-1.07	**4.50	-1.53	1.33	**23.24	**14.74	**40.98	**34.75
Cyp1b1	Cytochrome P450, family 1, subfamily b, polypeptide 1	-1.27	1.32	1.12	-1.33	**3.39	**3.32	**6.59	**4.92
Neurotrophin signalling	*3*	*Mmu04722*							

Gene symbol	Gene Name	Fold change							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Ntrk2	Neurotrophic tyrosine kinase	1.36	**2.27	1.24	1.26	*2.37	**2.81	*1.73	*1.49
Rps6ka2	Ribosomal protein s6 kinase, polypeptide 2	-1.09	1.05	-1.29	-1.26	**1.97	1.14	**1.81	*1.56
Small cell lung cancer	*3*	*Mmu05222*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Cdk6	Cyclin-dependent kinase 6	-1.1	*-1.56	1.22	-1.35	*-1.51	-1.22	**-1.92	**-1.98
Nfkbia	Nuclear factor of kappa light polypeptide geneenhancer in B-cells inhibitor, alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60
Nos3	Nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
Calcium signalling	*3*	*Mmu4020*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			

		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adrb2	Adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	*2.02	**2.12
Cckbr	Cholecystokinin B receptor	**-1.05	-1.06	-1.07	-1.05	**2.18	*1.50	*1.44	1.11
Nos3	Nitric oxide synthase 3, endothelial cell	1.12	1.22	-1.13	1.14	1.29	*1.85	**2.29	*1.79
Nod-like receptor signalling	*3*	*Mmu4621*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Ccl2	Chemokine (C-C motif) ligand 2	**2.39	*-1.81	**-2.25	-1.03	**-2.48	**-2.32	*-1.49	**-2.46
Ccl7	Chemokine (C-C motif) ligand 7	**3.22	*-2.18	**-3.06	-1.11	*-2.07	**-2.81	**-3.00	**-3.36
Nfkbia	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-alpha	-1.22	-1.19	*1.49	-1.05	*-1.37	*-1.45	*-1.45	**-1.60

Terpenoid backbone biosynthesis	*3*	*Mmu0380*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Hmgcr	3-hydroxy-3-methylglutaryl-coenzyme A reductase	1.07	-1.17	-1.02	-1.01	-1.31	*-1.70	-1.31	**-1.85
Hmgcs1	3-hydroxy-3-methylglutary-coenzyme A synthase	-1.27	*-2.00	1.01	-1.45	*-2.26	*-2.43	*-1.83	**-3.07
Idil	Isopentenyl-diphosphate delta isomerase	-1.57	-1.57	1.58	-1.36	-1.97	-2.15	-1.65	-3.86
Jak-stat	*2*	*Mmu04630*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Jak1	Janus kinase 1	-1.25	1.2	-1.05	-1.21	*1.97	**2.27	**2.08	*1.75
Spry1	Sprouty homolog 1	*1.92	*1.95	**-2.24	*-1.38	**2.46	*1.62	*2.30	*1.81
TGF-beta signalling	*2*	*Mmu04350*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8

Id1	Inhibitor of DNA binding 1	1.02	*-2.06	1.44	-1.29	*-3.46	*-3.93	*-2.02	**-2.71
Id3	Inhibitor of DNA binding 3	*-1.51	-1.08	*1.74	1.15	*-1.99	**-2.51	-1.21	-1.48
Neuroactive ligand-receptor interaction	*2*	*Mmu04080*							
Gene symbol	**Gene Name**	**Fold change**							
		Early time points				Late time points			
		D1T0	D2T0	D1T2	D2T2	D1T4	D2T4	D1T8	D2T8
Adrb2	Adrenergic receptor, beta 2	-1.17	*1.54	1.24	-1.13	**2.44	1.45	*2.02	**2.12
Cckbr	Cholecystokinin B receptor	**-1.05	-1.06	-1.07	-1.05	**2.18	*1.50	*1.44	1.11

* significant $p<0.05$
** significant FDR adjusted $p<0.05$
D1- Concentration of coal tar used for exposure = 1 μg/ml
D2- Concentration of coal tar used for exposure = 4 μg/ml
Tx- Hours of recovery time
(Example: D2T4, indicates that the sample was treated with a concentration of 4 μg/ml of coal tar and had a recovery time of 4 hours)
x.xx up-regulation
x.xx down-regulation

6 References

Abelev, G. I. (1971) Alpha-fetoprotein in ontogenesis and its association with malignant tumors. *Advances in Cancer Research,* **14,** 295.

Abidi, P., H. Zhang, S. M. Zaidi, W. J. Shen, S. Leers-Sucheta, Y. Cortez, J. Han & S. Azhar (2008) Oxidative stress-induced inhibition of adrenal steroidogenesis requires participation of p38 mitogen-activated protein kinase signaling pathway. *The Journal of Endocrinology,* **198,** 193.

Adachi, J., T. Fujita, R. Kudo, M. Asano, Y. Nurhantari & Y. Ueno (2003) 7-hydroperoxycholesterol and oxysterols as indices of oxidative stress: chronic ethanol feeding and rat skeletal muscle. *Legal Medicine (Tokyo, Japan),* **5 (1),** 105.

Adib, T. R., S. Henderson, C. Perrett, D. Hewitt, D. Bourmpoulia, J. Ledermann & C. Boshoff (2004) Predicting biomarkers for ovarian cancer using gene-expression microarrays. *British Journal of Cancer,* **90,** 686.

Agendia, MammaPrint http://row.agendia.com/en/about_mammaprint.html. Accessed March 2009.

Aihara, T., Fujiwara, Y., Ooka, M. Sakita, I. Tamaki, Y and Monden, M. (2004) Mammaglobin B as a novel marker for detection of breast cancer micrometastases in axillary lymph nodes by reverse transcription polymerase chain reaction *Breast Cancer Research and Treatment,* **58,** 137.

Aitio, A., J. Jarvisalo, M. Kiilunen, A. Tossavainen & P. Vaittinen (1984) Urinary excretion of chromium as an indicator of exposure to trivalent chromium sulphate in leather tanning. *International Archives of Occupational and Environmental Health,* **54,** 241.

Aly, H. A. & O. Domenech (2009) Cytotoxicity and mitochondrial dysfunction of 2,3,7,8-Tetrachlorodibenzo-dioxin (TCDD) in isolated rat hepatocytes. *Toxicology Letters,* **191,** 79. Epub 2009 Aug 15.

Andriole, G. L., E. D. Crawford, R. L. Grubb, 3rd, S. S. Buys, D. Chia, T. R. Church, M. N. Fouad, E. P. Gelmann, P. A. Kvale, D. J. Reding, J. L. Weissfeld, L. A. Yokochi, B. O'Brien, J. D. Clapp, J. M. Rathmell, T. L. Riley, R. B. Hayes, B. S. Kramer, G. Izmirlian, A. B. Miller, P. F. Pinsky, P. C. Prorok, J. K. Gohagan & C. D. Berg (2009) Mortality results from a randomized prostate-cancer screening trial. *The New England Journal of Medicine,* **360,** 1310.

ATSDR (2002) Toxicological profile for wood creosote, coal tar creosote, coal tar, coal tar pitch, and coal tar pitch volitiles. *U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry,* 2.

Autrup, H., B. Daneshvar, L. O. Dragsted, M. Gamborg, M. Hansen, S. Loft, H. Okkels, F. Nielsen, P. S. Nielsen, E. Raffn, H. Wallin & L. E. Knudsen (1999) Biomarkers for exposure to ambient air pollution--comparison of carcinogen-DNA adduct levels with other exposure markers and markers for oxidative stress. *Environmental Health Perspectives,* **107,** 233.

Autrup, H., A. B. Vestergaard & H. Okkels (1995) Transplacental transfer of environmental genotoxins: polycyclic aromatic hydrocarbon-albumin in non-smoking women, and the effect of maternal GSTM1 genotype. *Carcinogenesis,* **16,** 1305.

Balk, S. P., Y. J. Ko & G. J. Bubley (2003) Biology of prostate-specific antigen. *Journal Clinical Oncology,* **21,** 383.

Barraclough, C. A. & E. W. Haller (1970) Positive and negative feedback effects of estrogen on pituitary LH synthesis and release in normal and androgen-sterilized rats. *Endocrinology,* **86,** 542.

Bell, J. (1876) Paraffin epithelioma of the scrotum. *Edinburgh Medical Journal,* **22,** 135.

Belsito, D. V. & P. Kechijian (1982) The role of tar in Goeckerman therapy. *Archives of Dermatology,* **118,** 319.

Benjamini, Y. & Y. Hochberg (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* **57,** 289.

Benowitz, N. L. (1996) Cotinine as a biomarker of environmental tobacco smoke exposure. *Epidemiologic Reviews,* **18,** 188.

Benowitz, N. L. & P. Jacob, 3rd (1994) Metabolism of nicotine to cotinine studied by a dual stable isotope method. *Clinical Pharmacology and Therapeutics,* **56,** 483.

Berndt-Weis, M. L., L. M. Kauri, A. Williams, P. White, G. Douglas & C. Yauk (2009) Global transcriptional characterization of a mouse pulmonary epithelial cell line for use in genetic toxicology. *Toxicology In Vitro,* **23,** 816.

Bhatia, M., C. Landolfi, F. Basta, G. Bovi, R. D. Ramnath, A. C. de Joannon & A. Guglielmotti (2008) Treatment with bindarit, an inhibitor of MCP-1 synthesis, protects mice against trinitrobenzene sulfonic acid-induced colitis. *Inflammation Research,* **57,** 464.

Bhatia, M., R. D. Ramnath, L. Chevali & A. Guglielmotti (2005) Treatment with bindarit, a blocker of MCP-1 synthesis, protects mice against acute pancreatitis. *American Journal Physiology Gastrointestinal Liver Physiology,* **288,** g1259.

Bjorkhem, I. (2002) Do oxysterols control cholesterol homeostasis? *The Journal of Clinical Investigation,* **110,** 725.

Boaz, M., L. Iuliano, J. Himmelfarb, Z. Matas, F. Micheletta, E. McMonagle, V. Friedman, S. Natoli, G. Gvirtz, A. Biro, S. Smetana, G. Sabo, U. Gafter & T. Weinstein (2005) Baseline oxysterols and other markers of oxidative stress, inflammation and malnutrition in the vitamin e and intima media thickness progression in end-stage renal disease (VIPER) cohort. *Nephron,* **100,** c111.

Bonfoco, E., D. Krainc, M. Ankarcrona, P. Nicotera & S. A. Lipton (1995) Apoptosis and necrosis: two distinct events induced, respectively, by mild and intense insults with N-methyl-D-aspartate or nitric oxide/superoxide in cortical cell cultures. *Proceedings of the National Academy of Sciences of the United States of America,* **92,** 7162.

Brulet, P. & H. M. McConnell (1976) Lateral hapten mobility and immunochemistry of model membranes. *Proceedings of the National Academy of Sciences of the United States of America,* **73,** 2977.

Buntner, B., J. Sroczynski, M. Dwornicki & M. Snit (1991) [Effect of working conditions in the coke-processing department on the secretory function of the adrenal cortex]. *Medycyna Pracy,* **42,** 173.

Buterin, T., M. T. Hess, N. Luneva, N. E. Geacintov, S. Amin, H. Kroth, A. Seidel & H. Naegeli (2000) Unrepaired fjord region polycyclic aromatic hydrocarbon-DNA adducts in ras codon 61 mutational hot spots. *Cancer Research,* **60,** 1849.

Byfield, F. J., H. Aranda-Espinoza, V. G. Romanenko, G. H. Rothblat & I. Levitan (2004) Cholesterol depletion increases membrane stiffness of aortic endothelial cells. *Biophysical Journal,* **87,** 3336.

Calleros, L., M. Lasa, F. J. Rodriguez-Alvarez, M. J. Toro & A. Chiloeches (2006) RhoA and p38 MAPK mediate apoptosis induced by cellular cholesterol depletion. *Apoptosis,* **11,** 1161.

Campbell, I. K., M. J. Rich, R. J. Bischof & J. A. Hamilton (2000) The colony-stimulating factors and collagen-induced arthritis: exacerbation of disease by M-CSF and G-CSF and requirement for endogenous M-CSF. *Journal of Leukocyte Biology,* **68,** 144.

Carrell, C. J., H. L. Carrell, J. P. Glusker, E. Abu-Shaqara, C. Cortez & R. G. Harvey (1994) Bay- and fjord-region distortions in dibenz[a,j]anthracene and tetrabenzo[de,hi,mn,qr]naphthacene. *Carcinogenesis,* **15,** 2931.

Castano-Vinyals, G., A. D'Errico, N. Malats & M. Kogevinas (2004) Biomarkers of exposure to polycyclic aromatic hydrocarbons from environmental air pollution. *Occupational and Environmental Medicine,* **61,** e12.

CEPA, (1988) PSL1 Substances Considered as "Toxic" under Section 11 of the 1988 Canadian Environmental Protection Act

Chang, F., L. Wang, Q. Zhao, Q. Zhu, Y. Wu, C. Chen, S. Syrjanen & K. Syrjanen (1992) Investigation on the carcinogenic effects of coal tar pitch in rat respiratory tract by intratracheal instillations. *Carcinogenesis,* **13,** 233.

Chen, H., R. Vlahos, S. Bozinovski, J. Jones, G. P. Anderson & M. J. Morris (2005) Effect of short-term cigarette smoke exposure on body weight, appetite and brain neuropeptide Y in mice. *Neuropsychopharmacology,* **30,** 713.

Choi, K. W., C. Y. Lee, Y. K. Lee, Y. H. Kim & K. H. Kim (2009) The Photosensitivity Localized in a Vitiliginous Lesion Was Associated with the Intramuscular Injections of Synthetic Progesterone during an In Vitro Fertilization-embryo Transfer. *Annals of Dermatology,* **21,** 88.

Conti, M. E. (2008) *Biological Monitoring: Theory and Applications*. Billerica: WIT Press.

Costantino, J. P., C. K. Redmond & A. Bearden (1995) Occupationally related cancer risk among coke oven workers: 30 years of follow-up. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine,* **37,** 597.

Cui, X., J. T. Hwang, J. Qiu, N. J. Blades & G. A. Churchill (2005) Improved statistical tests for differential gene expression by shrinking variance components estimates. *Biostatistic,* **6,** 59.

Dennis, G., Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane & R. A. Lempicki (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology,* **4,** P3.

Ding, J., E. O. List, S. Okada & J. J. Kopchick (2009) Perspective: proteomic approach to detect biomarkers of human growth hormone. *Growth Hormone IGF Research,* **19,** 399.

DiSepio, D., R. A. Chandraratna & S. Nagpal (1999) Novel approaches for the treatment of psoriasis. *Drug Discovery Today,* **4,** 222.

Donnelly, K. C., L. D. Claxton, H. J. Huebner & J. L. Capizzi (1998) Mutagenic interactions of model chemical mixtures. *Chemosphere,* **37,** 1253.

Duffus, J. H. & H.G.J. Worth (2006) *Fundamental Toxicology*. Cambridge: RSC Publishing.

Dunn, J. M. & J.E.Van Eyk. (2008) Clinical proteomics: From Diagnosis to Therapy. Federal republic of Germany: Wiley-VCH Verglag GMBH & Co. KGaA Weinheim.

Feingold, K. R., M. Q. Man, G. K. Menon, S. S. Cho, B. E. Brown & P. M. Elias (1990) Cholesterol synthesis is required for cutaneous barrier function in mice. *The Journal of Clinical Investigation,* **86,** 1738.

Fenech, M. (1998) Important variables that influence base-line micronucleus frequency in cytokinesis-blocked lymphocytes-a biomarker for DNA damage in human populations. *Mutation Research,* **404,** 155.

Fleshner, N. E. & N. Lawrentschuk (2009) Risk of developing prostate cancer in the future: overview of prognostic biomarkers. *Urology,* **73,** S21.

Florescu, A., R. Ferrence, T. Einarson, P. Selby, O. Soldin & G. Koren (2009) Methods for quantification of exposure to cigarette smoking and environmental tobacco smoke: focus on developmental toxicology. *Therapeutic Drug Monitoring,* **31,** 14.

Goncalves, A., B. Esterni, F. Bertucci, R. Sauvan, C. Chabannon, M. Cubizolles, V. J. Bardou, G. Houvenaegel, J. Jacquemier, S. Granjeaud, X. Y. Meng, E. T. Fung, D. Birnbaum, D. Maraninchi, P. Viens & J. P. Borg (2006) Postoperative serum proteomic profiles may predict metastatic relapse in high-risk primary breast cancer patients receiving adjuvant chemotherapy. *Oncogene,* **25,** 981.

Gossen, J. A., W. J. de Leeuw, C. H. Tan, E. C. Zwarthoff, F. Berends, P. H. Lohman, D. L. Knook & J. Vijg (1989) Efficient rescue of integrated shuttle vectors from transgenic mice: a model for studying mutations in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* **86,** 7971.

Greenbaum, D., N. M. Luscombe, R. Jansen, J. Qian & M. Gerstein (2001) Interrelating different types of genomic data, from proteome to secretome: 'oming in on function. *Genome Research,* **11,** 1463.

Groopman, J. D., J. A. Hasler, L. J. Trudel, A. Pikul, P. R. Donahue & G. N. Wogan (1992) Molecular dosimetry in rat urine of aflatoxin-N7-guanine and other aflatoxin metabolites by multiple monoclonal antibody affinity chromatography and immunoaffinity/high performance liquid chromatography. *Cancer Research,* **52,** 267.

Grupper, C. (1971) The Chemistry, pharmacology and use of tar in treatment of psyoriasis, In: Farber EM, Cox AJ, eds : Stanford: Stanford University.

Guglielmotti, A., E. D'Onofrio, I. Coletta, L. Aquilini, C. Milanese & M. Pinza (2002) Amelioration of rat adjuvant arthritis by therapeutic treatment with bindarit, an inhibitor of MCP-1 and TNF-alpha production. *Inflammation Research,* **51,** 252.

Hamadeh, H. K., R. P. Amin, R. S. Paules & C. A. Afshari (2002a) An overview of toxicogenomics. *Current Issues in Molecular Biology,* **4,** 45.

Hamadeh, H. K., P. R. Bushel, S. Jayadev, K. Martin, O. DiSorbo, S. Sieber, L. Bennett, R. Tennant, R. Stoll, J. C. Barrett, K. Blanchard, R. S. Paules & C. A. Afshari (2002b) Gene expression analysis reveals chemical-specific profiles. *Toxicology Science,* **67,** 219.

Hanlon, P. R., W. Zheng, A. Y. Ko & C. R. Jefcoate (2005) Identification of novel TCDD-regulated genes by microarray analysis. *Toxicology and Applied Pharmacology,* **202,** 215.

Hecht, S. S., K. El-Bayoumy, A. Rivenson & S. Amin (1994) Potent mammary carcinogenicity in female CD rats of a fjord region diol-epoxide of benzo[c]phenanthrene compared to a bay region diol-epoxide of benzo[a]pyrene. *Cancer Research,* **54,** 21.

Heinrich-Ramm, R., R. Wegner, A. H. Garde & X. Baur (2002) Cotinine excretion (tobacco smoke biomarker) of smokers and non-smokers: comparison of GC/MS and RIA results. *International Journal of Hygiene and Environmental Health,* **205,** 493.

Henderson, R. F. (1995) Strategies for use of biological markers of exposure. *Toxicology Letters,* **82-83,** 379.

Heppell, S. A., N. D. Denslow, L. C. Folmar & C. V. Sullivan (1995) Universal assay of vitellogenin as a biomarker for environmental estrogens. *Environmental Health Perspectives,* **103 (7),** 9.

Hites, R. A. (1980) Environmental Carcinogens. *Science,* **207,** 1067.

Hong, J. Y. & C. S. Yang (1997) Genetic polymorphism of cytochrome P450 as a biomarker of susceptibility to environmental toxicity. *Environmental Health Perspectives,* **105 (4),** 759.

Huang da, W., B. T. Sherman & R. A. Lempicki (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols,* **4,** 44.

IARC (1983) Polynuclear aromatic compounds. International Agency for Research on Cancer *Polynuclear Aromatic Compounds, Part I, Chemical, Environmental, and experimental Data, World Health Organization, Lyon.*

Ichihara, S., Y. Yamada, G. Ichihara, T. Nakajima, P. Li, T. Kondo, F. J. Gonzalez & T. Murohara (2007) A role for the aryl hydrocarbon receptor in regulation of ischemia-induced angiogenesis. *Arteriosclerosis, thrombosis, and vascular biology,* **27,** 1297.

Ishii, K., J. Adachi, M. Tomita, M. Kurosaka & Y. Ueno (2002) Oxysterols as indices of oxidative stress in man after paraquat ingestion. *Free Radical Research,* **36,** 163.

Ivnitski-Steele, I. D., M. Friggens, M. Chavez & M. K. Walker (2005) 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) inhibition of coronary vasculogenesis is mediated, in part, by reduced responsiveness to endogenous angiogenic stimuli, including vascular endothelial growth factor A (VEGF-A). *Birth Defects Research,* **73,** 440.

Ivnitski-Steele, I. D. & M. K. Walker (2003) Vascular endothelial growth factor rescues 2,3,7,8-tetrachlorodibenzo-p-dioxin inhibition of coronary vasculogenesis. *Birth Defects Research,* **67,** 496.

Jauregui, H. O., N. T. Hayner, J. L. Driscoll, R. Williams-Holland, M. H. Lipsky & P. M. Galletti (1981) Trypan blue dye uptake and lactate dehydrogenase in adult rat hepatocytes--freshly isolated cells, cell suspensions, and primary monolayer cultures. *In Vitro,* **17,** 1100.

Jongeneelen, F. J., R. B. Anzion, P. T. Scheepers, R. P. Bos, P. T. Henderson, E. H. Nijenhuis, S. J. Veenstra, R. M. Brouns & A. Winkes (1988) 1-Hydroxypyrene in urine as a biological indicator of exposure to polycyclic aromatic hydrocarbons in several work environments. *The Annals of Occupational Hygiene,* **32,** 35.

Kandutsch, A. A. & H. W. Chen (1973) Inhibition of sterol synthesis in cultured mouse cells by 7alpha-hydroxycholesterol, 7beta-hydroxycholesterol, and 7-ketocholesterol. *The Journal of Biological Chemistry,* **248,** 8408.

Karakaya, A., S. Suzen, N. Vural & G. Oflaz (1993) Evaluation of the biological threshold value of urinary cadmium concentration in a group of workers. *Bulletin of Environmental Contamination and Toxicology,* **51,** 483.

Kerr, M. K. (2003) Design considerations for efficient and effective microarray studies. *Biometrics,* **59,** 822.

Kerr, M. K. & G. A. Churchill (2001) Statistical design and the analysis of gene expression microarray data. *Genetical Research,* **77,** 123.

Klassen (1996) *Casarett and Doull's Toxicology: The Basic Science of Poisons*. McGraw Hill Companies, Inc.

Klee, E. W., D. F. Carlson, S. C. Fahrenkrug, S. C. Ekker & L. B. Ellis (2004) Identifying secretomes in people, pufferfish and pigs. *Nucleic acids research,* **32,** 1414.

Krane, D. E., and Raymer, M.L. (2004) *Fundamental Concepts of Bioinformatic*. Elsevier Academic Press.

Krauss, M., W. Wilcke, C. Martius, A. G. Bandeira, M. V. Garcia & W. Amelung (2005) Atmospheric versus biological sources of polycyclic aromatic hydrocarbons (PAHs) in a tropical rain forest environment. *Environmental Pollution,* **135,** 143.

Legrand, C., J. M. Bour, C. Jacob, J. Capiaumont, A. Martial, A. Marc, M. Wudtke, G. Kretzmer, C. Demangel, D. Duval & et al. (1992) Lactate dehydrogenase (LDH) activity of the cultured eukaryotic cells as marker of the number of dead cells in the medium [corrected]. *Journal of Biotechnology,* **25,** 231.

Lemieux, C. L., (2006) Evaluating the mutagenic activities of Complex PAH Mixtures in Soil. MSc thesis: Biology. Carleton University, Ottawa.

Lemieux, C. L., I. B. Lambert, S. Lundstedt, M. Tysklind & P. A. White (2008) Mutagenic hazards of complex polycyclic aromatic hydrocarbon mixtures in contaminated soil. *Environmental toxicology and chemistry / SETAC,* **27,** 978.

MacAulay, E. J., (2008) TCDD Disrupts Gene Expression in the Developing Mouse Kidney. PhD: Department of Pharmaceutical Sciences. University of Toronto, Toronto.

Mahadevan, B., C. P. Marston, A. Luch, W. M. Dashwood, E. Brooks, C. Pereira, J. Doehmer & W. M. Baird (2007) Competitive inhibition of carcinogen-activating CYP1A1 and CYP1B1 enzymes by a standardized complex mixture of PAH extracted from coal tar. *International Journal of Cancer,* **120,** 1161.

Manahan (2005) *Environmental Chemistry,* 8th edition edn. Boca Raton, Florida: CRC Press.

Marek, K. & T. Wocka-Marek (1994) [Beta-n-acetylglucosaminidase in urine as a sign of kidney damage in workers exposed to metallic mercury]. *Medycyna pracy,* **45,** 101.

Marston, C. P., C. Pereira, J. Ferguson, K. Fischer, O. Hedstrom, W. M. Dashwood & W. M. Baird (2001) Effect of a complex environmental mixture from coal tar containing polycyclic aromatic hydrocarbons (PAH) on the tumor initiation, PAH-DNA binding and metabolic activation of carcinogenic PAH in mouse epidermis. *Carcinogenesis,* **22,** 1077.

Mathews, C. K., Holde, K.E. van, Ahern, Kevin, G. (2000) *Biochemestry,* 3rd Edition edn.: Robin Heyden.

Mbeunkui, F., O. Fodstad & L. K. Pannell (2006) Secretory protein enrichment and analysis: an optimized approach applied on cancer cell lines using 2D LC-MS/MS. *Journal of Proteome Research,* **5,** 899.

Mehrad, B., M. P. Keane & R. M. Strieter (2007) Chemokines as mediators of angiogenesis. *Thrombosis and Haemostasis,* **97,** 755.

Michaud, S. E., C. Menard, L. G. Guy, G. Gennaro & A. Rivard (2003) Inhibition of hypoxia-induced angiogenesis by cigarette smoke exposure: impairment of the HIF-1alpha/VEGF pathway. *The FASEB Journal,* **17,** 1150.

Mirolo, M., M. Fabbri, M. Sironi, A. Vecchi, A. Guglielmotti, G. Mangano, G. Biondi, M. Locati & A. Mantovani (2008) Impact of the anti-inflammatory agent bindarit on the chemokinome: selective inhibition of the monocyte chemotactic proteins. *European Cytokine Network,* **19,** 119.

Miyata, M., E. Tamura, K. Motoki, K. Nagata & Y. Yamazoe (2003) Thalidomide-induced suppression of embryo fibroblast proliferation requires CYP1A1-mediated activation. *Drug Metabolism and Disposition: The Biological Fate of Chemicals,* **31,** 469.

Moss, S. T. & J. A. Hamilton (2000) Proliferation of a subpopulation of human peripheral blood monocytes in the presence of colony stimulating factors may contribute to the inflammatory process in diseases such as rheumatoid arthritis. *Immunobiology,* **202,** 18.

Mukhtar, H., M. Das & D. R. Bickers (1986) Skin tumor initiating activity of therapeutic crude coal tar as compared to other polycyclic aromatic hydrocarbons in SENCAR mice. *Cancer Letters,* **31,** 147.

Muller, S. A. & H. O. Perry (1984) The Goeckerman treatment in psoriasis: six decades of experience at the Mayo Clinic. *Cutis; Cutaneous Medicine for the Practitioner,* **34,** 265.

National Institute of Standards and Technology, G., MD Certificate of Analysis: Standard Reference Material 1597a.

Needham, D. & R. S. Nunn (1990) Elastic deformation and failure of lipid bilayer membranes containing cholesterol. *Biophysical Journal,* **58,** 997.

Nielsen, P. S., H. Okkels, T. Sigsgaard, S. Kyrtopoulos & H. Autrup (1996) Exposure to urban and rural air pollution: DNA and protein adducts and effect of glutathione-S-transferase genotype on adduct levels. *International Archives of Occupational and Environmental Health,* **68,** 170.

NPRI, (2008). National Pollutant Release Inventory. http://www.ec.gc.ca/pdb/npri/consultations/PAH2005_e.cfm. Accessed March 2007

Okabe, T., R. Takayanagi, M. Adachi, K. Imasaki & H. Nawata (1998) Nur77, a member of the steroid receptor superfamily, antagonizes negative feedback of ACTH synthesis and secretion by glucocorticoid in pituitary corticotrope cells. *The Journal of Endocrinology,* **156,** 169.

Ovrevik, J., M. Lag, J. A. Holme, P. E. Schwarze & M. Refsnes (2009) Cytokine and chemokine expression patterns in lung epithelial cells exposed to components characteristic of particulate air pollution. *Toxicology,* **259,** 46.

Pan, G., T. Hanaoka, Y. Yamano, K. Hara, M. Ichiba, Y. Wang, J. Zhang, Y. Feng, Z. Shujuan, D. Guan, G. Gao, N. Liu & K. Takahashi (1998) A study of multiple biomarkers in coke oven workers--a cross-sectional study in China. *Carcinogenesis,* **19,** 1963.

Pathak, M. A. & P. C. Joshi (1983) The nature and molecular basis of cutaneous photosensitivity reactions to psoralens and coal tar. *The Journal of Investigative Dermatology,* **80 (l),** s66.

Pellitteri-Hahn, M. C., M. C. Warren, D. N. Didier, E. L. Winkler, S. P. Mirza, A. S. Greene & M. Olivier (2006) Improved mass spectrometric proteomic profiling of the secretome of rat vascular endothelial cells. *Journal of Proteome Research,* **5,** 2861.

Petrozzi, J. W., J. O. Barton, K. K. Kaidbey & A. M. Kligman (1978) Updating the Goeckerman regimen for psoriasis. *The British Journal of Dermatology,* **98,** 437.

Phillips, D. H. (1983) Fifty years of benzo(a)pyrene. *Nature,* **303,** 468.

Pittelkow, M. R., H. O. Perry, S. A. Muller, W. Z. Maughan & P. C. O'Brien (1981) Skin cancer in patients with psoriasis treated with coal tar. A 25-year follow-up study. *Archives of Dermatology,* **117,** 465.

Podechard, N., O. Fardel, M. Corolleur, M. Bernard & V. Lecureur (2009) Inhibition of human mesenchymal stem cell-derived adipogenesis by the environmental contaminant benzo(a)pyrene. *Toxicology In Vitro,* **23,** 1139.

Pohl, D., C. Andrys, L. Borska, Z. Fiala, K. Hamakova, K. Ettler & J. Krejsek (2009) CC and CXC chemokines patterns in psoriasis determined by protein array method were influenced by Goeckerman's therapy. *Acta Medica,* **52,** 9.

Pulfer, M. K. & R. C. Murphy (2004) Formation of biologically active oxysterols during ozonolysis of cholesterol present in lung surfactant. *The Journal of Biological Chemistry,* **279,** 26331.

Ramnath, R. D., S. W. Ng, A. Guglielmotti & M. Bhatia (2008) Role of MCP-1 in endotoxemia and sepsis. *International Immunopharmacology,* **8,** 810.

RDCT, (2005) R Developement Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roehm, N. W., G. H. Rodgers, S. M. Hatfield & A. L. Glasebrook (1991) An improved colorimetric assay for cell proliferation and viability utilizing the tetrazolium salt XTT. *Journal of Immunological Methods,* **142,** 257.

Roggi, C., C. Minoia, G. F. Sciarra, P. Apostoli, L. Maccarini, S. Magnaghi, A. Cenni, A. Fonte, G. F. Nidasio & G. Micoli (1997) Urinary 1-hydroxypyrene as a marker of exposure to pyrene: an epidemiological survey on a general population group. *The Science of the Total Environment,* **199,** 247.

Santodonato, J. (1997) Review of the estrogenic and antiestrogenic activity of polycyclic aromatic hydrocarbons: relationship to carcinogenicity. *Chemosphere,* **34,** 835.

Schena, M., D. Shalon, R. W. Davis & P. O. Brown (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science,* **270,** 467.

Schroepfer, G. J., Jr. (2000) Oxysterols: modulators of cholesterol metabolism and other processes. *Physiological Reviews,* **80,** 361.

Scott, A. S., Uhler, A.D., McCarthy, K.J., Emsbo-Mattingly, S. (2002) *Chemical Fingerprinting of Hydrocarbons*. New York: Elsevier Academic Press.

Searle, S. R., Speed, F.M., Milliken, G.A. (1980) The population marginal means in the linear model: An alternative to least squares means *The American Statistician* **34,** 216.

Shimada, T. (2006) Xenobiotic-metabolizing enzymes involved in activation and detoxification of carcinogenic polycyclic aromatic hydrocarbons. *Drug Metabolism and Pharmacokinetics,* **21,** 257.

Sigel, A., Sigel, H. & Sigel, R.K.O (2007) *The Ubiquitous Roles of Cytochrome450 Proteins. Metal Ions in Life Sciences*. Wiley; Chichester.

Sim, T. C., L. M. Reece, K. A. Hilsmeier, J. A. Grant & R. Alam (1995) Secretion of chemokines and other cytokines in allergen-induced nasal responses: inhibition by topical steroid treatment. *American Journal of Respiratory and Critical Care Medicine,* **152,** 927.

Strieter, R. M., T. J. Standiford, G. B. Huffnagle, L. M. Colletti, N. W. Lukacs & S. L. Kunkel (1996) "The good, the bad, and the ugly." The role of chemokines in models of human disease. *Journal Immunology,* **156,** 3583.

Suk, K. (2010) Combined analysis of the glia secretome and the CSF proteome: neuroinflammation and novel biomarkers. *Expert Review of Proteomics,* **7,** 263.

Swaen, G. M., J. J. Slangen, A. Volovics, R. B. Hayes, T. Scheffers & F. Sturmans (1991) Mortality of coke plant workers in The Netherlands. *British Journal of Industrial Medicine,* **48,** 130.

Swaen, G. M. & J. M. Slangen (1997) Mortality in a group of tar distillery workers and roofers. *International Archives of Occupational and Environmental Health,* **70,** 133.

Teoh, K. H., C. A. Bradley, J. Gauldie & H. Burrows (1995) Steroid inhibition of cytokine-mediated vasodilation after warm heart surgery. *Circulation,* **92,** 347.

Thami, G. P. & R. Sarkar (2002) Coal tar: past, present and future. *Clinical and Experimental Dermatology,* **27,** 99.

Thiermann, H., Szinicz, L., Eyer, P., Zilker, T., Worek, F. (2005) Correlation between red blood cell acetylcholinesterase activity and neuromuscular transmission in organophosphate poisoning. *Chemical and Biological Interaction,* **157,** 345.

van't Veer, L. J. & R. Bernards (2008) Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature,* **452,** 564.

van 't Veer, L. J., H. Dai, M. J. van de Vijver, Y. D. He, A. A. M. Hart, M. Mao, H. L. Peterse, K. van der Kooy, M. J. Marton, A. T. Witteveen, G. J. Schreiber, R. M. Kerkhoven, C. Roberts, P. S. Linsley, R. Bernards & S. H. Friend (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature,* **415,** 530.

Victor, F. C. & A. B. Gottlieb (2002) TNF-alpha and apoptosis: implications for the pathogenesis and treatment of psoriasis. *Journal of Drugs Dermatology,* **1,** 264.

Vondracek, J., P. Krcmar, J. Prochazkova, L. Trilecova, M. Gavelova, L. Skalova, B. Szotakova, M. Buncek, H. Radilova, A. Kozubik & M. Machala (2009) The role of aryl hydrocarbon receptor in regulation of enzymes involved in metabolic activation of polycyclic aromatic hydrocarbons in a model of rat liver progenitor cells. *Chemico-Biological Interactions,* **180,** 226.

White, P. A., G. R. Douglas, J. Gingerich, C. Parfett, P. Shwed, V. Seligy, L. Soper, L. Berndt, J. Bayley, S. Wagner, K. Pound & D. Blakey (2003) Development and characterization of a stable epithelial cell line from Muta Mouse lung. *Environmental and Molecular Mutagenesis,* **42,** 166.

Wilkins, L. W., (2006) **α**-fetoprotein protein In: *Stedman's Medical Dictionary.*

Wu, H., M. K. Kerr, X. Cui & G. A. Churchill (2003) *MAANOVA: A Software Package for the Analysis of Spotted cDNA Microarray Experiments.* Edited by: Parmigiani G, Garrett ES, Irizarry RA, Zeger S. Springer-Verlag.

Wu, J. T., L. Book & K. Sudar (1981) Serum alpha fetoprotein (AFP) levels in normal infants. *Pediatric Research,* **15,** 50.

Yang, W. M., Y. M. Zhao & J. J. Yang (1994) [Deposition of air particles with polycyclic aromatic hydrocarbons (PAH) in human respiratory tract]. *Zhonghua yu fang yi xue za zhi [Chinese Journal of Preventive Medicine],* **28,** 151.

Yang, Y. H., S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai & T. P. Speed (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic acids research,* **30,** e15.

Yeagle, P. L. (1985) Cholesterol and the cell membrane. *Biochimica et Biophysica Acta,* **822,** 267.

Zoja, C., D. Corna, G. Benedetti, M. Morigi, R. Donadelli, A. Guglielmotti, M. Pinza, T. Bertani & G. Remuzzi (1998) Bindarit retards renal disease and prolongs survival in murine lupus autoimmune disease. *Kidney International,* **53,** 726.