Citation for published version (APA):
Sub-chronic oral exposure to benzo(a)pyrene leads to distinct transcriptomic changes in the lungs that are related to carcinogenesis

Sarah Labib¹, Carole Yauk¹, Andrew Williams¹, Volker M Arlt², David H Phillips², Paul A White¹, Sabina Halappanavar¹*

¹ Environmental and Radiation Health Sciences Directorate, Health Canada, Ottawa, Ontario K1A 0K9, Canada
² Analytical and Environmental Sciences Division, King’s College London, London, SE1 9NH, United Kingdom

Sarah Labib, Sarah.Labib@hc-sc.gc.ca
Carole Yauk, Carole.Yauk@hc-sc.gc.ca
Andrew Williams, Andrew.Williams@hc-sc.gc.ca
Volker M Arlt, Volker.Arlt@kcl.ac.uk
David H. Phillips, David.Phillips@kcl.ac.uk
Paul A White, Paul.White@hc-sc.gc.ca
Sabina Halappanavar, Sabina.Halappanavar@hc-sc.gc.ca

* Corresponding Author: Sabina Halappanavar, Environmental and Radiation Health Sciences Directorate, Environmental Health Science and Research Bureau, Health Canada, Tunney’s Pasture, Bldg. 8 (P/L 0803A), 50 Colombine Driveway, Ottawa, Ontario K1A 0K9, Canada. Tel.: +1 613 957-3136; fax: +1 613 941 8530; sabina.halappanavar@hc-sc.gc.ca

Running Title – Transcriptomic response of BaP-exposed lungs
Abstract

We have previously shown that acute oral exposure to the environmental carcinogen benzo(a)pyrene (BaP) elicits comparable levels of DNA adducts, but distinct transcriptomic changes, in mouse lungs and livers, the two main BaP bioactivating organs. Oral BaP exposure is predominantly associated with lung cancer and not hepatic cancer in some animal models, suggesting that gene expression differences may provide insight into the drivers of tissue-specific carcinogenesis. In the present study, we examine pulmonary DNA adduct formation, lacZ mutant frequency, and mRNA profiles in adult male MutaTM Mouse following sub-chronic (28 day) oral exposure to BaP (0, 25, 50, and 75 mg/kg/day) and sacrificed 3 days post-exposure. The results are compared to those obtained from livers of the same mice (previously published). Although there was a 1.8- to 3.3-fold increase in the levels of DNA adducts in lung compared to liver, the lacZ transgene mutant frequency was similar in both tissues. At the transcriptomic level, a transition from activation of the DNA damage response p53 pathway at the low dose to the induction of genes involved in angiogenesis, evasion of apoptosis and growth signals at the high doses was evident only in the lungs. These results suggest that tissue DNA adducts and mutant frequency are sensitive markers of target tissue exposure and mode of action, whereas, early changes in gene expression may provide a better indication of the likelihood of carcinogenic transformation in selected tissues. Moreover, the study provides new information on the underlying mechanisms that contribute to tissue specific responses to BaP.
**Introduction**

Benzo(a)pyrene (BaP) is a well-studied polycyclic aromatic hydrocarbon (PAH) that is found in cigarette smoke, indoor and outdoor air, as well as in charred meats. Metabolic activation of BaP by cytochrome P450 (CYP) enzymes, namely CYP1a1 and CYP1b1, results in the generation of several BaP metabolites, including BaP-7,8-diol-9,10-epoxide (BPDE), a highly reactive form of BaP. BPDE is the primary metabolite of BaP that covalently binds DNA and forms DNA adducts (Baird et al., 2005). These DNA adducts, if left un repaired, can cause mutations in genes involved in xenobiotic metabolism, tumour suppression, or oncogenes leading to tumour development. The International Agency for Research on Cancer (IARC) has recently upgraded their classification of BaP to a Group 1 (i.e., known human) carcinogen (Baan et al., 2009). Exposure to BaP leads to tumour formation in animal models, and epidemiologically, a link between BaP exposure and cancer incidence in humans has been established (OEHHA, 2010). In addition to being carcinogenic and mutagenic, BaP is a known immunosuppressant (Dean et al., 1983). BaP can also alter cell cycle progression (Solhaug et al., 2005), induce inflammation (Qamar et al., 2012), and impair DNA repair and apoptotic processes (Solhaug et al., 2005) leading to aberrant cellular functioning.

The primary site of BaP metabolism is the liver; however, hepatocarcinogenesis is rare in humans and inconsistent in animals (OEHHA, 2010). Lung carcinogenesis, on the other hand, is prevalent in mice exposed orally to BaP (Stoner et al., 1984; Wattenberg and Leong, 1970), including Mutatm Mouse (Hakura et al., 1998). This is explained by greater BaP retention in the lung (Galván et al., 2005; Harrigan et al., 2004), higher induction of CYP1a1 and CYP1b1 in the lung tissue (Harrigan et al., 2006), and higher levels of BPDE-DNA adducts in the lungs compared to the liver (Harrigan et al., 2004). Indeed, the levels of BaP-induced DNA adducts in
any tissue at a given time is a function of metabolic conversion of BaP to its DNA-reactive metabolites via phase I xenobiotic metabolism, phase II detoxification, the rate of adduct repair, and the cell turnover rate, collectively suggesting that tissue dynamics and BaP response between lung and liver tissues likely differs.

Our previous work examined global gene expression changes in the lungs of adult male mice following acute exposure to 150 and 300 mg/kg BaP by oral gavage for three consecutive days and sacrificed four hours post-exposure (Halappanavar et al., 2011). The pulmonary transcriptomic response included changes in biological pathways involved in B-cell receptor (BCR) signalling, inflammation, and DNA damage response. We also noted that the response in the lung tissue was different, and in some respects, much more pronounced than livers from the same mice. Oxidative stress, xenobiotic metabolism, aryl hydrocarbon receptor (AhR) signalling, and glutathione metabolism were among the common pathways that were affected in both tissues, whereas negative regulation of the BCR signalling, indicative of potential primary immunodeficiency, was unique to the lung tissue. However, both tissues exhibited similar levels of DNA adducts, suggesting that evidence of higher levels of tissue DNA adducts are not sufficient to predict the observed transcriptomic differences in lungs and livers. Although the study was conducted at very high doses of BaP, the findings support the hypothesis that tissue-specific mechanisms are at work, and that differential regulation of the known biological pathways associated with cancer formation could explain the selective targeting of lungs for carcinogenesis.

In the present study, we explore further the potential of early BaP-induced gene expression changes to identify the tissue-specific events that may eventually lead to carcinogenic transformation following BaP exposure. We globally profile the pulmonary transcriptome and
measure in parallel the induced frequency of mutations and DNA adducts (known predictors of genotoxicity-mediated cancer) in the lung tissues of adult male mice repeatedly exposed to 25, 50 and 75 mg/kg BaP for 28 consecutive days by oral gavage and sacrificed three days post-exposure. The results are compared and interpreted in light of the differential biological responses reported for the lungs and livers of mice following both acute (Halappanavar et al., 2011; Yauk et al., 2011) and sub-chronic (Malik et al., 2012) exposures to BaP.
Materials and Methods

Animal Treatment

Adult (25 weeks-old) male Muta\textsuperscript{TM}Mouse (transgenic mouse strain 40.6) were exposed to BaP (Sigma Aldrich, Canada) as described previously (Malik et al., 2012). In brief, animals were dosed daily via oral gavage for 28 days with varying doses of BaP (0, 25, 50, 75 mg/kg body weight/day) dissolved in olive oil. Each dose group contained five animals. Mice were sacrificed by cardiac puncture under isofluorane anaesthesia 72 hours following the final exposure. The right lobe of the lung was excised, flash frozen in liquid nitrogen, and stored at $-80^\circ$C until use. For the duration of the experiment, food (2014 Teklad Global standard rodent diet) and water were provided \textit{ad libitum} and mice were caged individually in plastic film isolators (Harlan Isotec, UK) on a 12 hour light/12 hour dark cycle. Mice were bred, maintained and treated in accordance with the Canadian Council for Animal Care Guidelines and approved by Health Canada’s Animal Care Committee.

Tissue Selection

The major exposures to BaP occur via the oral route (drinking and feed) (Hettemer-Frey and Travis, 1991). The organs directly affected by BaP as a consequence of oral exposure include stomach, esophagus, tongue, and larynx (Culp \textit{et al.}, 1998). However, studies conducted by Stoner et al. (1984) and Wattenberg and Leong (1970) show lung and liver to be equally impacted by BaP (oral gavage) and moreover, revealed sensitivity of lung for tumour development in comparison with liver. In alignment with these reports, our previous work (Halappanavar \textit{et al.}, 2011; Yauk \textit{et al.}, 2011) revealed lung-specific regulation of the biological processes known to be associated with cancer formation. Thus, in the present study we have investigated transcriptional responses in lung and liver. The results of the study may also help
address the importance of considering the multi organ toxicity in calculating the risk associated with chemicals, such as BaP.

**Tissue DNA Extraction**

The frozen lung and liver tissue was sliced randomly. Genomic DNA was isolated from a random tissue section for measuring the levels of DNA adducts and transgene mutant frequency. In brief, lung tissue was minced and de-gassed to remove all traces of air in the alveoli. Livers were homogenized in ice cold TMST buffer (50 mM Tris, pH 7.6, 3 mM magnesium acetate, 250 mM sucrose, 0.2% Triton X-100) as described in Douglas et al. (1994). The minced tissue was washed twice in cold phosphate buffered saline and lysed in 10 mM Tris pH 7.6, 10 µM EDTA, 100 µM NaCl, 1% SDS overnight at 37°C on a rotating platform. The lysate was digested with proteinase K (1 mg/ml lysis buffer). DNA was isolated using a serial phenol/chloroform/isoamyl alcohol (25:24:1) and chloroform/isoamyl alcohol (24:1) extraction (Renault et al., 1997). DNA was precipitated in ethanol and dissolved in Tris-EDTA buffer (10 mM Tris pH 7.6, 500 mM EDTA). DNA was stored at 4°C until use.

**DNA Adduct Analysis**

DNA adduct formation in each sample was determined using the nuclease P1 digestion enrichment version of the $^{32}$P-postlabelling assay as described previously (Phillips and Arlt, 2007) with minor modifications. Briefly, 4 µg of DNA was digested overnight with micrococcal nuclease (288 mUnits, Sigma, cat. no. N3755) and calf spleen phosphodiesterase (1.2 mUnits, MP Biomedicals, cat. no. 100977), enriched and labelled as described elsewhere (Phillips and Arlt, 2007). Radiolabelled adducted nucleotide biphosphates were separated by thin-layer chromatography on polyethyleneimine-cellulose plates (Macherey-Nagel, Düren, Germany) with the following chromatographic conditions (Arlt et al., 2008): D1, 1.0 M sodium phosphate, pH 6;
D3, 4.0 M lithium-formate, 7.0 M urea, pH 3.5; D4, 0.8 M LiCl, 0.5 M Tris, 8.5 M urea, pH 8. Chromatographs were scanned using a Packard Instant Imager (Dowers Grove, USA) and DNA adduct levels (relative adduct labelling) were calculated from the adduct counts per minute (cpm), the specific activity of $[^\gamma-^{32}P]ATP$, and the amount of DNA (pmol of DNA-P) used. An external BPDE-DNA standard was used for identification of BaP-DNA adducts. Results are expressed as DNA adducts/10$^8$ nucleotides.

**LacZ Mutant Frequency (Positive Selection)**

The transgene lacZ mutant frequency in lungs was determined using the P-gal (phenyl-β-D-galactopyranoside) positive selection assay as described in Vijg and Douglas (1996) and Lambert et al. (2005). The λgt10lacZ DNA was rescued from the genomic DNA using the Transpack$^{TM}$ lambda packaging system (Stratagene, USA). The packaged phage particles were mixed with host bacterium (*Escherichia coli* lacZ-, galE-, recA-, pAA119 with galT and galK), plated on minimal medium containing 0.3% (w/v) P-Gal and incubated overnight at 37$^0$C. Total plaque-forming units (pfu) were measured on concurrent titres that did not contain P-Gal. Mutant frequency is expressed as the ratio of the number of mutant pfu to total pfu. Mutant frequency data analysis was performed as described previously (Malik et al., 2012).

**RNA Extraction and Purification**

Total RNA was isolated for gene expression analysis and qRT-PCR validation as described previously (Halappanavar et al., 2011). Briefly, total RNA was extracted from the lungs using TRIzol reagent (Invitrogen, USA) and purified using RNeasy Mini Kit (Qiagen, Canada). The RNA quantity and purity was checked using a Nanodrop Spectrophotometer (Thermo Fisher Scientific, Canada). The RNA integrity was determined using an Agilent 2100
Bioanalyzer (Agilent Technologies, Canada). The samples showing $A_{260}/A_{280}$ ratios between 2.1 and 2.2 and having RNA Integrity Number above 7.5 were used for further analysis.

**Microarray Hybridization and Analysis**

Double-stranded cDNA and cyanine-labelled cRNA was synthesized (Agilent Linear Amplification Kits, Agilent Technologies, Canada) from 250 ng of total RNA from each sample and universal reference total RNA (Stratagene, Canada). Cyanine-labelled cRNA targets were *in vitro* transcribed using T7 RNA polymerase and purified by RNeasy Mini Kit (Qiagen, Canada). From each (sample and reference) labelled sample 825 ng of cRNA was hybridized to Agilent 4X44K oligonucleotide microarrays (Agilent Technologies, Canada) at 60°C overnight (16 hours) in the Agilent SureHyb hybridization chamber. Arrays were washed and scanned on an Agilent G2505B Scanner according to manufacturer’s recommendations. Feature extraction software version 10.7.3.1 (Agilent Technologies, Canada) was used to extract the data.

A reference design (Kerr and Churchill, 2001) was used to analyse mRNA expression as described previously (Malik et al., 2012). All analyses were conducted in the R (R-Development-Core-Team, 2010) environment using the MAANOVA library (Wu, 2010). The background fluorescence was measured using the negative control (-)3xSLv1 probes; probes with median signal intensities less than the trimmed mean (trim = 5%) plus three trimmed standard deviations of (-)3xSLv1 probe were flagged as absent (within the background signal). Probes were considered present if at least four of the five samples within a condition had signal intensities greater than three trimmed standard deviations above the trimmed mean of the (-)3xSLv1 probes (background signal). Data were normalized using the transform.madata() function using the glowess option with a span of 0.1. Ratio intensity plots and heat maps for the raw and normalized data were constructed to identify outliers. One sample (50 mg/kg group) was
removed from the analysis based on clustering. The statistical model for this analysis included fixed effects of array and treatment condition and was applied to the log_2 of the absolute intensities. Differentially expressed transcripts (upregulated or downregulated relative to olive oil treated control mouse lung samples) were determined using the Fs statistic option in the matest() function. The p values for all statistical tests were estimated by the permutation method with residual shuffling and false discovery rate (FDR) adjusted p values were estimated using the adjPval() function. The fold change calculations were estimated as described previously (Malik et al., 2012). Significant genes were selected based on a FDR adjusted p value < 0.05 for any BaP exposed versus control contrast.

qRT-PCR Array Validation

Mouse pathway-specific PCR array (cancer-PAMM-033, SABiosciences™, USA) and custom PCR arrays consisting of 172 genes in total were employed to validate the microarray results. Genes for the custom array were selected based on their implication in biological processes relevant to lung carcinogenesis. These genes included statistically significant differentially expressed genes (FDR adjusted p < 0.05), differentially expressed genes that exhibited high fold changes (fold rank only) but were not statistically significant, and genes that were differentially regulated (FDR adjusted p < 0.05) in the livers from the same mice (Malik et al., 2012). In brief, 0.8 µg of total lung tissue RNA (n=5/group) from each sample was reverse transcribed using RT2 First Strand Kit (SABiosciences™, USA). Real-time PCR was performed using RT2 SYBR Green PCR Master Mix on a CFX96 real-time detection system (Bio-Rad, Canada). Threshold cycle values for each well were averaged. Relative gene expression was determined according to the comparative Ct method and normalized to reference RNAs Hprt1, Gapdh, and Actb housekeeping genes for the Cancer Pathway array and to reference RNAs Gusb
and Hprt1 housekeeping genes for the custom arrays. Transcripts were further normalized by subtracting the median delta Ct value for each sample. Differential expression was determined with a two-sample bootstrap test using R software (R-Development-Core-Team, 2010). The fold change was estimated using the ratio of the arithmetic mean of the treated sample to the mean of the control samples. Standard errors for the fold change values were estimated using the bootstrap test (Efron and Tibshirani, 1993).

**Bioinformatics**

All mRNA data are deposited in the NCBI Gene Expression Omnibus database under accession numbers GSE35718 (lung) and GSE24910 (liver). Following normalization, biological functions perturbed in response to BaP were identified using functional annotation clustering in the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang da et al., 2009) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000). The biological and molecular functions of genes that were significantly differentially expressed (FDR adjusted $p \leq 0.05$, or FDR adjusted $p \leq 0.1$ and fold change $\geq 1.5$) following exposure to BaP treatment were analysed and explored in IPA (Ingenuity® Systems, Redwood City, CA). Molecular relational networks of genes modulated by BaP in lung tissue enriched for cancer function were generated using IPA. Each molecule was overlaid onto a global molecular network developed from information contained in the Ingenuity Knowledge Base. Networks were generated based on their connectivity. All relationships are supported by at least one reference from the literature.
Results

Daily exposure to 25, 50, and 75 mg/kg BaP for 28 consecutive days caused no overt signs of toxicity and there was no significant body weight loss in any of the exposed mice compared to vehicle treated controls.

BaP-induced DNA adduct formation in lung and liver tissue

The presence of DNA adducts indicates the development of pre-mutagenic lesions following a chemical exposure. The $^{32}$P-postlabelling technique was employed to measure BaP-induced DNA adducts in the lung tissues of mice exposed to 0, 25, 50, and 75 mg/kg BaP for 28 consecutive days (Figure 1a). BaP-7,8-diol-9,10-epoxide-$N^2$-deoxyguanosine (dG-$N^2$-BPDE) was identified as the main adduct as reported previously (Lemieux et al., 2011). No adducts were detected in mice dosed with vehicle control. DNA adduct formation by BaP was dose-dependent. Although liver DNA adducts were analyzed in these samples previously (Lemieux et al., 2011; Malik et al., 2012), these data were not used. Instead, $^{32}$P post-labelling was repeated in these samples in order to ensure that the protocols in both liver and lung were identical and performed in the same lab. Similar to results observed in the lungs, dG-$N^2$-BPDE was the main adduct identified in liver, and adduct levels increased in a dose-dependent manner (Figure 1a). However, the frequency of DNA adducts in liver was significantly lower than in lung tissue for each dose (Bonferonni-corrected $p$ value $\leq 0.05$), suggesting that lung tissue is more susceptible to DNA damage caused by BaP.

Mutagenic activity of BaP in the lung tissue

The Muta$^\text{TM}$Mouse contains around 29 +/- 4 (Shwed et al., 2010) copies of $\lambda$gt10lacZ shuttle vector stably integrated in the mouse genome, thus permitting in vivo lacZ mutant
frequency analysis. In agreement with the DNA adduct results, a dose-dependent increase in mutant frequency was observed (Figure 1b) in lung tissues. Compared to the vehicle controls, a 5.2-, 13- and 18.6-fold increase was seen in the 25, 50 and 75 mg/kg dose groups, respectively. Mice gavaged with pure olive oil (vehicle) yielded a low spontaneous lacZ mutant frequency. No significant differences between the lung and liver tissues were observed for lacZ mutant frequency.

**General overview of pulmonary gene expression profiles**

Repeated oral gavage of mice with BaP induced significant changes in the expression of many genes in the lung tissue. MAANOVA analysis revealed 20, 145, and 373 unique probes that were differentially expressed (up or downregulated) with a fold change $\geq 1.5$ in either direction, and an FDR adjusted $p \leq 0.05$ in the 25, 50 and 75 mg/kg exposure groups compared to controls, respectively (Supplementary File 1). The complete lung microarray dataset is available through the Gene Expression Omnibus at NCBI (http://www.ncbi.nlm.nih.gov/geo/), accession number GSE35718. Hierarchical cluster analysis on all differentially expressed genes (FDR $p \leq 0.05$, fold change $\pm 1.5$) revealed that all treatment groups clustered separately from the control (Supplementary File 2), thus, a clear treatment effect was observed as a result of exposure to BaP.

Among the differentially expressed genes, 297 were upregulated and 90 were downregulated. Cyclin-dependent kinase inhibitor 1A ($Cdkn1a$), growth differentiation factor 15 ($Gdf15$), carboxylesterase 5 ($Ces5$), pleckstrin homology-like domain, family A, member 3 ($Phlda3$), multimerin 2 ($Mmrn2$), cyclin G1 ($Ccng1$), serum amyloid A 3 ($Saa3$) and sestrin 2 ($Sesn2$) were the most upregulated genes in all the dose groups. The most downregulated genes
included Sarcolipin (Sln), myomesin 2 (Myom2), tropomyosin 1, alpha (Tpm1α) in the 75 mg/kg group, chitinase 3-like 3 (Chi3l2) and gastrin releasing peptide (Grp) in the 50 mg/kg group, and sineoculis-related homeobox 1 homolog (Six1) (Drosophila) and ras homolog gene family, member U (Rhou) in the 25 mg/kg dose group. Twenty-five probes representing 23 genes were commonly altered between the dose groups with an FDR \( p \leq 0.05 \) in at least one dose group (Table 1). The observed transcriptomic response was dose-dependent.

Gene ontology analysis was employed to assign biological processes and functional categories to significant genes (FDR \( p \leq 0.05 \), fold change \( \geq 1.5 \)) in DAVID. Biological functions perturbed in response to BaP were identified using functional annotation clustering in DAVID (Supplementary File 3a). Significant functional enrichment (score \( > 2.0 \)) was found for biological processes including extracellular signalling, cell adhesion, angiogenesis, apoptosis, and regulation of cell growth in all dose groups. Further combined analysis employing DAVID and KEGG pathways revealed that the altered genes are associated with specific biological pathways, including p53 signalling, molecular mechanisms of cancer, small cell lung cancer, extracellular membrane receptor interaction, and focal adhesion (Supplementary File 4a).

In addition to the analysis described above, we relaxed the FDR adjusted \( p \) value to include genes that exhibit FDR \( p \leq 0.1 \) and fold change \( \geq 1.5 \) to see if we had missed any important genes in the pathways described above as a consequence of the stringent MAANOVA analysis. Using this criterion, 47, 181, and 429 genes in the 25, 50, and 75 mg/kg dose groups, respectively, were differentially altered. Fifty-one additional genes were identified by relaxing the \( p \)-value cut-off (Supplementary File 6). DAVID analysis (Supplementary File 3b and 4b) as described above revealed three new functional categories: cell-to-cell and cell-to-matrix interactions, protein maturation, and inflammatory response. An in-depth functional analysis
using IPA on these genes (FDR $p \leq 0.1$ and fold change $\geq 1.5$) revealed that the altered genes were associated with multiple functions, such as cancer, cell death, cell cycle, cellular growth and proliferation, cellular movement, DNA replication and repair, inflammatory response, and tumour morphology (Figure 2a). Since the present work is focused on finding perturbations in pathways associated with carcinogenic transformation in the lung tissue, a sub-analysis on specific genes that were grouped under the functional group ‘cancer’ was performed. The sub-analysis revealed that a majority of the genes from this group were mainly associated with the p53 signalling pathway and were implicated in several biological processes regulated by p53, such as cell cycle arrest, apoptosis, inhibition of cellular growth and proliferation, and DNA repair. Many of these genes were common across the dose groups (Tables 1, 2, and 3) and showed a clear dose-response. The other affected pathways included the molecular mechanisms of cancer, AhR signalling, BCR signalling, and angiogenesis (Table 3). A similar functional analysis using IPA on liver (Malik et al., 2012) revealed that the altered genes were associated with lipid metabolism, cancer, molecular transport, inflammatory response, and cell death (Figure 2b). Sub-analysis of the cancer-associated genes in the liver reveals that in addition to the p53 signalling pathway, this category includes genes involved in metabolism and inflammation, which are indirectly linked to carcinogenesis.

Visualization of the relationship between these genes in a biological network created in IPA (Figure 3) revealed that at the low dose the immediate downstream effectors of p53, such as Cdkn1a, Ccng1, Bax, and Tp53inp1 mainly involved in DNA damage recognition, cell cycle arrest, and apoptosis were upregulated. However, with increasing dose, genes known to be the suppressors of p53 signalling pathway, such as Mdm2 were upregulated along with other genes
implicated in proliferation and growth, evasion of apoptosis, promotion of angiogenesis, and cell invasion and metastasis pathways, all of which are the hallmarks of cancer.

**RT-PCR validation of microarray results**

Exposure to BaP resulted in altered expression of a number of genes involved in the molecular mechanisms of cancer (Table 3). Perturbation of this pathway was confirmed using a cancer pathway-specific PCR array (Mouse Cancer Pathway Finder PCR array, SABiosciences™) containing 84 genes. In addition, 61 more genes implicated in BCR signalling, calcium signalling, TGFβ signalling, cell cycle, and DNA damage repair pathways were also analysed (Table 3). From each dose group five individual samples were analysed.

In total, 57 of the 145 genes were differentially expressed with \( p \leq 0.1 \) in at least one treatment group compared to controls (Table 3, Supplementary File 5); 23 of these correlated with the microarray results. Thirty-one genes were significantly differentially expressed as measured by the PCR arrays but not by microarrays. This is presumably due to the enhanced dynamic range and sensitivity of RT-PCR. These genes included baculoviral IAP repeat-containing 5 (Birc5), tumour necrosis factor receptor superfamily member 1A (Tnfrsf1a), jun proto-oncogene (Jun), c-fos-induced growth factor (Figf), insulin-like growth factor 1 (Igf1), interferon beta 1 fibroblast (Ifnb1), vascular endothelial growth factors a and c (Vegfa and Vegfc), plasminogen activator urokinase (Plau), integrin alpha 3 (Itga3), platelet-derived growth factors a and b (Pdgfa and Pdgfb), and transforming growth factor beta 1 (Tgfb1) (Table 3).
Discussion

BaP is the only PAH classified by IARC as a known human carcinogen (Baan et al., 2009). The ability to induce DNA damage and mutations in key genes involved in tumour suppression or tumour promotion is one of the postulated mechanisms of BaP-induced carcinogenesis (Denissenko et al., 1996; Ross and Nesnow, 1999). The metabolic activation of BaP by phase I CYP enzymes is a prerequisite for its potentially carcinogenic interaction with DNA (Arlt et al., 2008; Rubin, 2001). Several clinical and epidemiological studies have shown a link between the levels of DNA adducts and exposure-induced carcinogenesis (Gunter et al., 2007; Tang et al., 2001). Thus, the presence of DNA adducts may be a risk factor for developing cancer. Although many tissues are capable of metabolizing BaP to its carcinogenic intermediate, which is associated with the development of mutations, carcinogenesis is specific to certain tissues. For example, BaP is predominantly metabolized in the liver; however, hepatocarcinogenesis is relatively less frequent in some experimental animal models and in humans compared to pulmonary carcinogenesis (Hakura et al., 1998; Stoner et al., 1984; Wattenberg and Leong, 1970). Several factors are suggested to play a vital role in the tissue-specific transformation process, including a balance between the metabolic activation and detoxification processes and the relative rates of DNA damage, DNA repair and individual susceptibility.

The presence of tissue DNA adducts and genetic mutations in genes is often used as a measure of the carcinogenic potential of a chemical (Mei et al., 2006; Wang et al., 2011). Although these markers of exposure and effect, respectively, classify a chemical as genotoxic and non-genotoxic, they may not accurately predict with certainty their carcinogenic potential. We have recently shown that acute exposure to BaP by oral gavage leads to similar levels of DNA adducts in the lungs and livers, two BaP bioactivating organs (Halappanavar et al., 2011).
However, despite similar levels of DNA damage (reflective of bioavailable BaP) in both the tissues, the transcriptomic responses in lungs were different and included lung-specific perturbations of several biological pathways associated with molecular mechanisms of cancer (Halappanavar et al., 2011). The present study was undertaken to further explore the molecular mechanisms underlying such distinct tissue-specific responses and their implication in cancer formation employing a sub-chronic exposure model.

In agreement with the published literature (Baird et al., 2005), we found that daily exposure to BaP for 28 days resulted in significant genotoxicity as measured by a dose-dependant increase in DNA adduct formation and lacZ transgene mutant frequency in the lungs of exposed animals (Figure 1). Compared to the livers from the same mice there was a statistically significant 1.8- to 3.3-fold increase (Bonferonni-corrected $p$ value $\leq 0.05$) in DNA adduct formation in the lung tissues. In alignment with these results, mRNA expression of DNA polymerase Kappa ($PolK$), necessary for the error-free bypass of the (+)-trans-BPDE-$N^2$-dG DNA adduct (Ogi et al., 2002), was significantly upregulated in the lungs (Table 3) but not the liver. In contrast, levels of DNA adducts in liver and lung were comparable in C57BL/6 mice acutely exposed to very high doses of BaP (Halappanavar et al., 2011) (i.e., 150 mg/kg and 300 mg/kg BaP per day for three days, sampled 4 and 24 hours after the last exposure). The discrepancy in DNA adduct levels between the acute and sub-chronic repeated dose study could be simply be due to the amount of total BaP gavaged over time. Compared to the hepatic tissue, rate of clearance of BaP from the lungs occurs slowly resulting in longer retention of BaP in lungs (Harrigan et al., 2004), suggesting that the repeated exposure to BaP over 28 consecutive days may have led to accumulation of BaP in lungs resulting in higher DNA adduct levels. Studies have shown accumulation of DNA adducts in some organs that are resistant to repair or following chronic repeated exposures. It has
also been suggested that at higher doses metabolic activation/detoxification of BaP could reach a threshold, which in turn may limit the speed with which reactive metabolites are detoxified in the organ compared to the speed of the translocation of these metabolites from the primary site (e.g., liver) to other distant organs (e.g., lungs).

Although we report significant differences in DNA adduct formation between the lung and liver, the transgene mutant frequency in both tissues was not significantly different, which may suggest more efficient DNA repair in the lungs compared to livers (Fig 2, high dose). The elevated mutant frequency observed in livers could be attributed to the higher rates of cellular metabolism (as reflected by the number of differentially altered genes in metabolic pathways) potentially resulting in an increase in by-products of reactive oxygen species. These results suggest that in addition to mutations arising as a result of unrepaired bulky DNA adducts, other factors such as alterations in critical molecular processes implicated in cancer formation may play a role in the selective targeting of the lung tissue for carcinogenesis following BaP exposure.

Our previous work (Halappanavar et al., 2011) on mice acutely exposed to BaP revealed suppression of B-cell receptor (BCR) signalling pathway and an elaborated DNA damage response in the lungs, whereas the livers from the same mice showed subtle changes in only a few genes associated with DNA damage response (Halappanavar et al., 2011). Repressed BCR signalling, leading to primary immunodeficiency, is potentially implicated in cancer initiation (Jumaa et al., 2005). In the present study we observed similar suppression of a few of the key genes in the BCR signalling pathway in the lungs. However, the response was only observed in the high dose group (Table 2), suggesting that the collective downregulation of the BCR
signalling pathway may be a potential mechanism that operates in the lungs following exposure to high doses of BaP over a very short period of time.

Exposure to BaP for 28 consecutive days induced a large transcriptomic response in the lungs sampled 3 days post-exposure compared to the livers. MAANOVA analysis of liver tissue revealed differential expression of 6, 7, and 121 unique probes compared to 20, 145, and 373 probes in the lungs, in the 25, 50 and 75 mg/kg exposure groups respectively (Malik et al., 2012). The p53 signalling pathway was commonly affected in both tissues (Table 2 and 5) (Malik et al., 2012). P53 is a tumour suppressor protein and mutations in the TP53 gene have been found in many types of human cancers linked to environmental exposures (Kucab et al., 2010) including lung cancers. The p53 protein is a transcription factor that mediates expression of several genes, including Cdkn1a, an inhibitor of cyclin-dependent kinases, Bax, a pro-apoptotic gene, Gadd45, a DNA damage repair gene, which is also involved in cell cycle arrest, and many inflammatory modulators and growth factors. Transgenic mice devoid of TP53 gene (p53 knockout) or mice in which p53 activity is chemically inhibited, exhibit increased susceptibility to develop spontaneous and chemically induced tumours (Donehower et al., 1992). Cellular DNA damage response to BaP is regulated by p53 (Park et al., 2006). Furthermore, increased rates of lung cancer following exposure to BaP arise in transgenic mice expressing lung-specific dominant-negative mutant p53 (Tchou-Wong et al., 2002), suggesting an important role for the p53 signalling pathway in the development of BaP-induced lung cancer.

Since the focus of the present study was to understand the molecular perturbations in lungs that are associated with tissue-specific tumour development (i.e, relative to the liver), we compared the expression changes of several genes in KEGG pathways associated with cancer in both tissues. Table 4 shows a list of 41 genes whose expression was significantly differentially
expressed (FDR $p \leq 0.05$, fold change $\geq 1.5$) in lungs but not in the livers. These genes are known for their function in biological pathways such as apoptosis, cell cycle, cytokine-cytokine receptor interaction, extracellular membrane receptor interactions, focal adhesion, mitogen-activated protein kinase (MAPK) signalling, p53 signalling, peroxisome proliferating agent receptor (PPAR) signalling, transforming growth factor beta (TGFβ) signalling, and Wnt signalling. In contrast, the liver showed expression changes in fewer genes, and these were mainly the downstream effectors of p53 signalling; *Apaf1*, *Trp63*, *MyoD*, *Egr1*, and *Btg2* involved in induction of apoptosis and negative regulation of cellular proliferation. Distinctly affected in the lungs were Sestrin 2 (*Sesn2*), *Pmaip1*, and *Trp53inp1*, associated with the p53-mediated DNA damage response, *Mdm2*, involved in the negative regulation of p53 signalling, and *Bcl2*, implicated in cell survival.

Carcinogenic transformation in the lung following BaP exposure appears to be a dose-dependent process. Figure 3 is a molecular relational network consisting of differentially altered genes in lung tissues enriched for cancer function (using IPA functional categorization) in response to BaP exposure in our model. Interestingly, nearly all genes in the network were found directly or indirectly connected to the transcription factor p53, which is the main node in the network. At the lowest dose, genes associated with an early response to DNA damage, such as *Cdkn1a* and *Ccng1* involved in cell cycle arrest and *Bax* and *Tp53inp1* involved in the apoptotic process, were upregulated, indicating cellular efforts to halt replication until the damage is repaired. With increasing dose, the p53 DNA damage response pathway continued to be activated. However, activation of pathways and genes implicated in silencing p53 function also became evident. This silencing may be related to AhR-dependent induction of *Mdm2* expression (Pääjärvi *et al.*, 2005), which is a negative regulator of p53 signalling; *Mdm2* was upregulated in
the two higher doses. While an apparent dose-dependent transition in the p53 DNA damage response pathway was observed, at the highest dose many genes involved in angiogenesis (Angpt2, Bdnf, Cyr61), evasion of apoptosis (Bcl2l1, Cryab, Lcn2, Rela), growth signals (Pdgfb, Pdgfd, Itga1, Notch4) and invasion and metastasis (Il1b, Mmp3, Pdgfra, Shh) were upregulated. These high dose effects are indicative of a shift in the fine cellular dynamics from a protective repair mode to a cellular replication and proliferation mode, which may suggest cellular transformation, potentially a key event in the progression to lung carcinogenesis.

One limitation of the present study is that the responses were investigated in whole tissue. The observed changes reflect the responses of over 40 different cell types. It is clear from the lung DNA adduct and mutant frequency data that the observed changes in molecular pathways are the direct results of BaP metabolism in the lung tissue. However, it should be noted that the influence of circulating factors from other target sites may also contribute to the pulmonary responses.

In conclusion, we show that the BaP-induced DNA adduct levels in lung tissue are comparatively higher than in livers after 28 consecutive days of exposure to increasing doses of BaP by oral gavage. However, the occurrence of fixed mutations was comparable in both tissues suggesting that for certain carcinogens traditional markers of genotoxicity, i.e., tissue DNA adduct levels and mutant frequency may only predict the exposure to the carcinogen and not the carcinogenic outcome. In contrast, dose-dependent transition in p53 DNA-damage response pathway that included tissue protective (adaptive) responses at the lowest dose (for e.g., genes involved in cell cycle and apoptosis) to growth and proliferative responses at the highest dose (e.g., growth signals and pro-proliferation) was distinct to lungs only, suggesting such key
biological pathway perturbations occurring early after the exposure may be causal to carcinogenesis in lungs compared to livers.
**SUPPLEMENTARY FILE DESCRIPTIONS**

**Supplementary File 1. Significant gene list.**

List of all significantly differentially expressed genes in at least 1 treatment group (FDR $p \leq 0.05$, fold change $\geq 1.5$) in response to repeated oral exposure to 25, 50, and 75 mg/kg/day BaP. The list is sorted from highest to lowest fold change in the 75 mg/kg/day treatment group.

**Supplementary File 2. Hierarchical cluster analysis.**

Hierarchical cluster analysis of all samples on differentially expressed genes. The heat map represents all genes that were statistically significant and differentially expressed in any of the treatment groups relative to control mice. Green bars represent low expression levels and red represent high expression levels relative to the reference channel. Black bars are similar to the reference channel. The heat map was generated using GeneSpring (Agilent Technologies). Clusters are color-coded: controls red, 25 mg/kg blue, 50 mg/kg brown, 75 mg/kg grey.

**Supplementary File 3. Gene ontology analysis.**

List of the most enriched functional annotation clusters using DAVID based on significantly differentially expressed genes in at least one dose with an FDR adjusted $p \leq 0.05$ (a) or FDR adjusted $p \leq 0.1$ (b) and fold change $\geq 1.5$. Each cluster groups similar annotations together to clarify the biological function. The names given to each annotation were arbitrarily chosen based on the constituents of the cluster. The Enrichment Score is the geometric mean of the $p$ values of the cluster constituents transformed by -log. The $p$ value is the EASE score which is a modified Fisher Exact $p$ value.
Supplementary File 4. Pathway analysis.

Functional annotation into KEGG pathway categories using significantly differentially expressed genes with an FDR adjusted $p$ value $\leq 0.05$ (a) or FDR adjusted $p$ value $\leq 0.1$ (b) and fold change $\geq 1.5$ in at least one dose group. The pathway annotations displayed have an EASE score (one-tail Fisher Exact $p$ value) less than 0.1.

Supplementary File 5. PCR validation.

All genes validated by qRT PCR using PCR arrays in lungs from BaP exposed mice. Bold genes are validated from microarrays. Genes colored in grey had low expression in both the dose groups and control (Ct>30).

Supplementary File 6. Relaxed gene list.

List of genes additionally differentially expressed following relaxing of the significance cut-off from FDR $p < 0.05$ to FDR $p < 0.1$. $P$ values shown are FDR adjusted.
**FUNDING INFORMATION**

This work was supported by Health Canada’s Genomics Research and Development Initiative and the Health Canada Chemicals Management Plan. Work at King’s College London was supported by Cancer Research UK.

**ACKNOWLEDGEMENTS**

We thank Byron Kuo for bioinformatics support, Lynda Soper for P-gal positive selection assay and Christine Lemieux for organizing and conducting the animal exposures.
REFERENCES


OEHHA (2010). Benzo(a)pyrene. Public health goals for chemicals in drinking water. In, Eds.).


FIGURE LEGENDS

Figure 1
DNA adduct formation (a) and lacZ mutant frequency (b) in the lungs and livers from Muta™ Mouse sub-chronically exposed to BaP. Levels of BaP-7,8-diol-9,10-epoxide-N²-deoxyguanosine (dG-N²-BPDE) adducts were determined using the nuclease P1 enrichment version of the ³²P post-labelling method. Data are represented as average ± standard error of the mean (n = 5 mice/group). ND, not detected. Average lacZ mutant frequency was determined using the P-Gal positive selection assay. Values shown are average frequencies x 10⁵ ± standard error of the mean. Asterisk (*) indicates significance by Student’s t-test (p < 0.01) compared to controls. Transgene mutant frequency data for the liver is from the study by (Malik et al., 2012).

Figure 2
Functional analysis of differentially expressed genes (FDR adjusted p ≤ 0.1 and fold change ≥ 1.5 in either direction) using IPA for lung (a) and liver (b). The Ingenuity Pathway Analysis Knowledge Base was used to identify biological functions and/or diseases that were significantly over-represented among the differentially expressed genes. A right-tailed Fisher’s exact test was used to calculate a p value determining the probability that each biological function and/or disease assigned to that data set was due to chance alone (threshold set at p value < 0.05).

Figure 3
Molecular relational network of genes modulated by BaP in lung tissues enriched for cancer function using IPA for all three doses with p53 at the central node. A network is a graphical
representation of the molecular relationships between molecules. Molecules are represented as nodes and the biological relationship between two nodes is represented as a line. The gene list was derived from cancer-associated genes determined in the Ingenuity Pathway Analysis functional analysis of significantly differentially expressed genes (Figure 2). The networks were generated from cancer-associated genes differentially regulated in the 25 mg/kg/day (green circle), 50 mg/kg/day (blue circle), and 75 mg/kg/day (pink circle) dose groups. Red colour denotes upregulation and green colour denotes downregulation.
<table>
<thead>
<tr>
<th>Agilent Probe</th>
<th>ACCESSION #</th>
<th>GENE SYMBOL</th>
<th>Function</th>
<th>25 mg/kg/day</th>
<th>50 mg/kg/day</th>
<th>75 mg/kg/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_52_P612803</td>
<td>NM_009831</td>
<td>Ccng1</td>
<td>Cell Cycle</td>
<td>↑ 1.8</td>
<td>↑ 3.3</td>
<td>↑ 4.6</td>
</tr>
<tr>
<td>A_51_P363947</td>
<td>NM_007669</td>
<td>Cdkn1a</td>
<td>Cell Cycle</td>
<td>↑ 2.5</td>
<td>↑ 5.7</td>
<td>↑ 9.3</td>
</tr>
<tr>
<td>A_51_P472751</td>
<td>NM_007915</td>
<td>Ei24</td>
<td>Apoptosis</td>
<td>↑ 1.4</td>
<td>↑ 1.8</td>
<td>↑ 2.1</td>
</tr>
<tr>
<td>A_51_P508289</td>
<td>NM_010145</td>
<td>Ephpx1</td>
<td>Xenobiotic Metabolism</td>
<td>↑ 1.7</td>
<td>↑ 2.3</td>
<td>↑ 2.9</td>
</tr>
<tr>
<td>A_52_P532982</td>
<td>NM_011819</td>
<td>Gdf15</td>
<td>Cellular Differentiation and Proliferation</td>
<td>↑ 2.6</td>
<td>↑ 6.9</td>
<td>↑ 9.1</td>
</tr>
<tr>
<td>A_51_P468173</td>
<td>NM_008251</td>
<td>Hmgn1</td>
<td>DNA Damage Repair</td>
<td>↑ 1.3</td>
<td>↑ 1.5</td>
<td>↑ 1.5</td>
</tr>
<tr>
<td>A_51_P167489</td>
<td>XM_140451</td>
<td>Lama3</td>
<td>Cell Adhesion</td>
<td>↑ 1.6</td>
<td>↑ 1.5</td>
<td>↑ 1.5</td>
</tr>
<tr>
<td>A_51_P484111</td>
<td>NM_016762</td>
<td>Matn2</td>
<td>Unknown Function</td>
<td>↓ 1.5</td>
<td>↓ 1.5</td>
<td>↓ 1.9</td>
</tr>
<tr>
<td>A_51_P463562</td>
<td>NM_008620</td>
<td>Mpa2</td>
<td>Possibly Cellular Differentiation</td>
<td>↑ 1.4</td>
<td>↑ 1.5</td>
<td>↑ 1.8</td>
</tr>
<tr>
<td>A_52_P13389</td>
<td>NM_026793</td>
<td>Myct1</td>
<td>Possibly Apoptosis</td>
<td>↑ 1.6</td>
<td>↑ 1.8</td>
<td>↑ 1.9</td>
</tr>
<tr>
<td>A_51_P195875</td>
<td>NM_010929</td>
<td>Notch4</td>
<td>Cellular Differentiation</td>
<td>↑ 1.6</td>
<td>↑ 1.7</td>
<td>↑ 1.8</td>
</tr>
<tr>
<td>A_51_P363801</td>
<td>NM_023217</td>
<td>Pgpep1</td>
<td>Proteolysis</td>
<td>↑ 1.3</td>
<td>↑ 1.6</td>
<td>↑ 1.5</td>
</tr>
<tr>
<td>Accession</td>
<td>Gene ID</td>
<td>Description</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------------------------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>A_51_P329928</td>
<td>NM_013750</td>
<td>Phlda3 Apoptosis</td>
<td>↑ 2.1</td>
<td>↑ 3.8</td>
<td>↑ 5.1</td>
<td></td>
</tr>
<tr>
<td>A_51_P224564</td>
<td>NM_176833</td>
<td>Ppm1f Apoptosis</td>
<td>↑ 1.4</td>
<td>↑ 1.5</td>
<td>↑ 2.0</td>
<td></td>
</tr>
<tr>
<td>A_51_P227392</td>
<td>NM_133955</td>
<td>Rhou Cell Cycle</td>
<td>↓ 1.6</td>
<td>↓ 1.6</td>
<td>↓ 1.5</td>
<td></td>
</tr>
<tr>
<td>A_51_P246903</td>
<td>NM_026467</td>
<td>Rps27l Apoptosis/DNA Damage Repair</td>
<td>↑ 1.4</td>
<td>↑ 1.8</td>
<td>↑ 2.5</td>
<td></td>
</tr>
<tr>
<td>A_51_P319572</td>
<td>NM_001037709</td>
<td>Rusc2 Unknown Function</td>
<td>↑ 1.3</td>
<td>↑ 1.5</td>
<td>↑ 1.7</td>
<td></td>
</tr>
<tr>
<td>A_51_P415755</td>
<td>AK052232</td>
<td>Sema6d Cellular Differentiation</td>
<td>↑ 1.4</td>
<td>↑ 1.4</td>
<td>↑ 1.5</td>
<td></td>
</tr>
<tr>
<td>A_52_P154710</td>
<td>AK170547</td>
<td>Sesn2 DNA Damage Repair</td>
<td>↑ 1.7</td>
<td>↑ 2.7</td>
<td>↑ 3.8</td>
<td></td>
</tr>
<tr>
<td>A_52_P220810</td>
<td>NM_144551</td>
<td>Trib2 Cellular Proliferation</td>
<td>↑ 1.4</td>
<td>↑ 1.9</td>
<td>↑ 2.1</td>
<td></td>
</tr>
<tr>
<td>A_51_P175580</td>
<td>NM_021897</td>
<td>Trp53inp1 Apoptosis</td>
<td>↑ 1.6</td>
<td>↑ 2.5</td>
<td>↑ 2.9</td>
<td></td>
</tr>
<tr>
<td>A_52_P503387</td>
<td>NM_021897</td>
<td>Trp53inp1 Apoptosis</td>
<td>↑ 1.6</td>
<td>↑ 2.3</td>
<td>↑ 2.9</td>
<td></td>
</tr>
<tr>
<td>A_51_P509609</td>
<td>NM_053082</td>
<td>Tspan4 Cell Signaling</td>
<td>↑ 1.7</td>
<td>↑ 1.7</td>
<td>↑ 1.8</td>
<td></td>
</tr>
<tr>
<td>A_51_P155052</td>
<td>NM_053247</td>
<td>Xlkdl Metastasis</td>
<td>↑ 2.2</td>
<td>↑ 2.3</td>
<td>↑ 3.2</td>
<td></td>
</tr>
<tr>
<td>A_52_P686785</td>
<td>NM_053247</td>
<td>Xlkdl Metastasis</td>
<td>↑ 2.1</td>
<td>↑ 2.1</td>
<td>↑ 3.6</td>
<td></td>
</tr>
</tbody>
</table>
Table 2

P53 regulated genes that were significantly differentially expressed with microarrays and real-time PCR in lungs. Only genes with FDR adjusted $p \leq 0.1$ and fold change ≥ 1.5 are shown. Gene symbols in bold were also differentially expressed in liver (Malik et al., 2012). ↑ indicates upregulated, - indicates no change.

<table>
<thead>
<tr>
<th>Gene</th>
<th>25 mg/kg/day</th>
<th>50 mg/kg/day</th>
<th>75 mg/kg/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>Microarray</td>
<td>RT-PCR</td>
<td>Microarray</td>
</tr>
<tr>
<td>Cdkn1a</td>
<td>↑ 2.5</td>
<td>↑ 3.2</td>
<td>↑ 5.7</td>
</tr>
<tr>
<td>Ccng1</td>
<td>↑ 1.7</td>
<td>↑ 2.2</td>
<td>↑ 3.1</td>
</tr>
<tr>
<td>Trp53inp1</td>
<td>↑ 1.6</td>
<td>↑ 1.8</td>
<td>↑ 2.4</td>
</tr>
<tr>
<td>Sesn2</td>
<td>-</td>
<td>↑ 2.9</td>
<td>↑ 2.2</td>
</tr>
<tr>
<td>Bax</td>
<td>↑ 1.5</td>
<td>↑ 1.9</td>
<td>↑ 1.9</td>
</tr>
<tr>
<td>Pmaip1</td>
<td>-</td>
<td>-</td>
<td>↑ 1.7</td>
</tr>
<tr>
<td>Mdm2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bcl2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CCnb2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tnfrsf10b</td>
<td>-</td>
<td>↑ 2.0</td>
<td>-</td>
</tr>
<tr>
<td>Zmat3</td>
<td>-</td>
<td>↑ 1.8</td>
<td>↑ 1.8</td>
</tr>
</tbody>
</table>
Table 3

Genes validated using real-time PCR arrays divided based on functional association in lungs. Only genes with FDR adjusted $p \leq 0.1$ and fold change $\geq 1.5$ are shown. ↑ indicates upregulated, ↓ indicates downregulated, - indicates no change. Genes with an asterix (*) indicate genes significantly differentially expressed by the PCR array but not significant by microarray. Please note that some genes appear more than once because they are implicated in several functions.

<table>
<thead>
<tr>
<th>Function</th>
<th>Gene Symbol</th>
<th>25 mg/kg/day</th>
<th>50 mg/kg/day</th>
<th>75 mg/kg/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Microarray</td>
<td>RT-PCR</td>
<td>Microarray</td>
</tr>
<tr>
<td>P53 Signalling</td>
<td>Cdkn1a</td>
<td>↑ 2.5</td>
<td>↑ 3.2</td>
<td>↑ 5.7</td>
</tr>
<tr>
<td></td>
<td>Tnfrsf10b*</td>
<td>-</td>
<td>↑ 2.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Ccng1</td>
<td>↑ 1.7</td>
<td>↑ 2.2</td>
<td>↑ 3.1</td>
</tr>
<tr>
<td></td>
<td>Bax</td>
<td>↑ 1.5</td>
<td>↑ 1.9</td>
<td>↑ 1.9</td>
</tr>
<tr>
<td></td>
<td>Sesn2</td>
<td>-</td>
<td>↑ 2.9</td>
<td>↑ 2.2</td>
</tr>
<tr>
<td></td>
<td>Trp53inp1</td>
<td>↑ 1.6</td>
<td>↑ 1.8</td>
<td>↑ 2.4</td>
</tr>
<tr>
<td></td>
<td>Mdm2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Zmat3</td>
<td>-</td>
<td>↑ 1.8</td>
<td>↑ 1.8</td>
</tr>
<tr>
<td></td>
<td>Ccnb2*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Pmaip1</td>
<td>-</td>
<td>-</td>
<td>↑ 1.7</td>
</tr>
<tr>
<td></td>
<td>Bcl2*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pathway</td>
<td>Gene</td>
<td>Expression 1</td>
<td>Expression 2</td>
<td>Expression 3</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>B cell Receptor Signalling</strong></td>
<td>Cd3g*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Cxcr5*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Dock2*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tnfsf10</td>
<td>-</td>
<td>↑ 1.5</td>
<td>↑ 1.8</td>
</tr>
<tr>
<td><strong>Calcium Signalling</strong></td>
<td>Penk*</td>
<td>-</td>
<td>↑ 2.5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Cyr6l</td>
<td>-</td>
<td>↑ 1.5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Adrb1</td>
<td>-</td>
<td>↑ 1.7</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>S100a8*</td>
<td>-</td>
<td>-</td>
<td>↑ 1.5</td>
</tr>
<tr>
<td></td>
<td>Slc30a1</td>
<td>-</td>
<td>-</td>
<td>↑ 1.5</td>
</tr>
<tr>
<td></td>
<td>Pln*</td>
<td>-</td>
<td>↓ 2.4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Per1*</td>
<td>-</td>
<td>↑ 2.4</td>
<td>-</td>
</tr>
<tr>
<td><strong>Transforming Growth Factor Signalling</strong></td>
<td>Ccnb2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Ctgf</td>
<td>-</td>
<td>↑ 1.7</td>
<td>-</td>
</tr>
<tr>
<td><strong>Transforming Growth Factor Signalling</strong></td>
<td>Cyr6l</td>
<td>-</td>
<td>↑ 1.5</td>
<td>-</td>
</tr>
<tr>
<td><strong>Transforming Growth Factor Signalling</strong></td>
<td>Eng*</td>
<td>-</td>
<td>↑ 1.8</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Jun*</td>
<td>-</td>
<td>↑ 1.5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Igf1*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Tgfb1*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gene</td>
<td>( Cdkn1a )</td>
<td>( Sesn2 )</td>
<td>( Mdm2 )</td>
<td>( Mgmt^* )</td>
</tr>
<tr>
<td>------------</td>
<td>---------------</td>
<td>----------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>↑ 2.5</td>
<td>↑ 3.2</td>
<td>↑ 5.7</td>
<td>↑ 8.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell Cycle and DNA Damage Repair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑ 2.9</td>
<td>↑ 2.2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ 1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ 2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ 1.7</td>
<td>↑ 1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ 1.6</td>
<td>↑ 1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ 1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apoptosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑ 1.5</td>
<td>↑ 1.9</td>
<td>↑ 1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑ 1.6</td>
<td>↑ 1.8</td>
<td>↑ 2.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ 1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tnfrsf10b*</td>
<td></td>
<td>↑ 2.0</td>
<td>-</td>
<td>↑ 2.7</td>
</tr>
<tr>
<td>Bax</td>
<td>↑ 1.5</td>
<td>↑ 1.9</td>
<td>↑ 1.9</td>
<td>↑ 2.5</td>
</tr>
<tr>
<td>Trp53inp1</td>
<td>↑ 1.6</td>
<td>↑ 1.8</td>
<td>↑ 2.4</td>
<td>↑ 2.1</td>
</tr>
<tr>
<td>Bcl2l1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>↑ 1.7</td>
</tr>
<tr>
<td>Tnfrsf1a*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bcl2*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tgfb1*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tnfsf10</td>
<td>↑ 1.5</td>
<td>↑ 1.5</td>
<td>↑ 1.8</td>
<td>-</td>
</tr>
<tr>
<td>Cidea*</td>
<td>-</td>
<td>↑ 2.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Card6*</td>
<td>-</td>
<td>↑ 1.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gene</td>
<td>Angiogenesis</td>
<td>Invasion and Metastasis</td>
<td>Adhesion</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>-------------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Coll18a1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Cyr61</td>
<td>- ↑ 1.5</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Kdr*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Pdgfa*</td>
<td>- ↑ 1.7</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Vegfc*</td>
<td>- ↑ 1.6</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Pecam1*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Igf1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Pdgfb</td>
<td>- ↑ 1.5</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Tgfβ1*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Vegfa*</td>
<td>- ↑ 1.5</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>F2*</td>
<td>- ↑ 2.9</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Plg*</td>
<td>- ↑ 1.9</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Serpine1</td>
<td>- ↑ 1.8</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Plau*</td>
<td>- ↑ 2.8</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Mmp2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Twist1*</td>
<td>- ↓ 2.9</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Pecam1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Itga3*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Signal Transduction</td>
<td><em>Jun</em></td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-------</td>
<td>-----</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><em>Myc</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4

Cancer pathway (KEGG) associated genes significantly differentially expressed in microarrays in the lung and not in the liver. Only genes with FDR adjusted $p \leq 0.05$ and fold change $\geq 1.5$ are shown. ↑ indicates upregulated, ↓ indicated downregulated, - indicates no change.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Symbol</th>
<th>25 mg/kg/day</th>
<th>50 mg/kg/day</th>
<th>75 mg/kg/day</th>
<th>Gene</th>
<th>Symbol</th>
<th>25 mg/kg/day</th>
<th>50 mg/kg/day</th>
<th>75 mg/kg/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bax</td>
<td>-</td>
<td>↑ 1.9</td>
<td>↑ 2.5</td>
<td></td>
<td>Map4k4</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.8</td>
</tr>
<tr>
<td>Bcl2l1</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.6</td>
<td>Mdm2</td>
<td>-</td>
<td>↑ 1.7</td>
<td></td>
<td>↑ 2.2</td>
</tr>
<tr>
<td>Bdnf</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.5</td>
<td>Myc</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 1.6</td>
</tr>
<tr>
<td>Bmp6</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.5</td>
<td>Nkd1</td>
<td>-</td>
<td>↑ 1.7</td>
<td></td>
<td>↑ 1.9</td>
</tr>
<tr>
<td>Ccl5</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 2.1</td>
<td>Ntrk2</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 1.6</td>
</tr>
<tr>
<td>Cd47</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 1.5</td>
<td>Pdgfd</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.5</td>
</tr>
<tr>
<td>Col4a1</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.7</td>
<td>Pdgfd</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.5</td>
</tr>
<tr>
<td>Col4a2</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.7</td>
<td>Pdgfra</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 1.5</td>
</tr>
<tr>
<td>Cpt1b</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 1.6</td>
<td>Pmaip1</td>
<td>-</td>
<td>↑ 1.7</td>
<td></td>
<td>↑ 1.8</td>
</tr>
<tr>
<td>Cpt1c</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.9</td>
<td>Ptprm</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.6</td>
</tr>
<tr>
<td>Dcn</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 2.4</td>
<td>Rasgrp3</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.8</td>
</tr>
<tr>
<td>Ei24</td>
<td>-</td>
<td>↑ 1.8</td>
<td>↑ 2.1</td>
<td></td>
<td>relA</td>
<td>-</td>
<td></td>
<td></td>
<td>↑ 1.6</td>
</tr>
<tr>
<td>Flnc</td>
<td>-</td>
<td></td>
<td></td>
<td>↓ 1.7</td>
<td>Serpine1</td>
<td>-</td>
<td>↑ 2.0</td>
<td></td>
<td>↑ 2.8</td>
</tr>
<tr>
<td>Gene</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gp9</td>
<td>-</td>
<td>↑ 1.8</td>
<td>↑ 1.5</td>
<td>Sesn2</td>
<td>↑ 1.7</td>
<td>↑ 2.7</td>
<td>↑ 3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itga1</td>
<td>-</td>
<td>-</td>
<td>↑ 1.5</td>
<td>Smurf2</td>
<td>-</td>
<td>↑ 1.6</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitol</td>
<td>-</td>
<td>↑ 1.6</td>
<td>↑ 1.5</td>
<td>Tnfrsf12a</td>
<td>-</td>
<td>-</td>
<td>↑ 1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lama3</td>
<td>↑ 1.6</td>
<td>↑ 1.5</td>
<td>↑ 1.5</td>
<td>Tnfsf10</td>
<td>-</td>
<td>↑ 1.8</td>
<td>↑ 1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lamb3</td>
<td>↑ 1.5</td>
<td>↑ 1.5</td>
<td>↑ 1.5</td>
<td>Ubc</td>
<td>-</td>
<td>-</td>
<td>↑ 1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lmac2</td>
<td>-</td>
<td>-</td>
<td>↑ 1.5</td>
<td>Wnt10a</td>
<td>-</td>
<td>↓ 1.6</td>
<td>↓ 1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ltb</td>
<td>-</td>
<td>-</td>
<td>↓ 1.5</td>
<td>Wnt7a</td>
<td>-</td>
<td>-</td>
<td>↑ 1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map4k1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5

P53 regulated genes identified in liver by Malik et al. (2012) were cross validated in the lung using real-time PCR arrays. Only genes with FDR adjusted $p \leq 0.1$ and fold change $\geq 1.5$ are shown. ↑ indicates upregulated, ↓ indicates downregulated, - indicates no change.

<table>
<thead>
<tr>
<th>Gene</th>
<th>25 mg/kg/day</th>
<th>50 mg/kg/day</th>
<th>75 mg/kg/day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lung</td>
<td>Liver</td>
<td>Lung</td>
</tr>
<tr>
<td></td>
<td>Symbol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdkn1a</td>
<td>↑ 3.2</td>
<td>-</td>
<td>↑ 8.1</td>
</tr>
<tr>
<td>Tnfrsf10b</td>
<td>↑ 2.0</td>
<td>-</td>
<td>↑ 2.7</td>
</tr>
<tr>
<td>Egr</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Btg2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ccnb2</td>
<td>-</td>
<td>-</td>
<td>↑ 1.7</td>
</tr>
<tr>
<td>Zmat3</td>
<td>↑ 1.8</td>
<td>-</td>
<td>↑ 1.7</td>
</tr>
<tr>
<td>Apaf1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Prc1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ccnb1</td>
<td>↑ 2.2</td>
<td>-</td>
<td>↑ 3.0</td>
</tr>
<tr>
<td>Birc5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Myod</td>
<td>-</td>
<td>↓ 2.1</td>
<td>-</td>
</tr>
<tr>
<td>Bax</td>
<td>↑ 1.9</td>
<td>-</td>
<td>↑ 2.5</td>
</tr>
<tr>
<td>Trp63</td>
<td>-</td>
<td>↓ 2.1</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 1

(A) DNA Adducts per 10^8 nucleotides

(B) Average Mutant Frequency (x10^-5)

Legend:
- Liver
- Lung
Figure 2