Published in final edited form as:

Toxicology. 2015 October 02; 336: 99–112. doi:10.1016/j.tox.2015.07.005.

Screening a mouse liver gene expression compendium identifies modulators of the aryl hydrocarbon receptor (AhR)

Keiyu Oshida^{a,1}, Naresh Vasani^a, Russell S. Thomas^{b,2}, Dawn Applegate^c, Frank J. Gonzalez^d, Lauren M. Aleksunes^e, Curtis D. Klaassen^f, J. Christopher Corton^{a,*}

^aNHEERL, US-EPA, Research Triangle Park, NC 27711, United States

^bThe Hamner Institutes for Health Sciences, Research Triangle Park, NC 27709, United States

^cRegeneMed, San Diego, CA 92121, United States

^dNational Cancer Institute, NIH, Bethesda, MD 20892, United States

^eRutgers University, Piscataway, NJ 08854, United States

^fUniversity of Kansas Medical Center, Kansas City, KS 66160, United States

Abstract

The aryl hydrocarbon receptor (AhR) is a ligand-activated transcription factor that mediates the biological and toxic effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), dioxin-like compounds (DLC) as well as some drugs and endogenous tryptophan metabolites. Short-term activation of AhR can lead to hepatocellular steatosis, and chronic activation can lead to liver cancer in mice and rats. Analytical approaches were developed to identify biosets in a genomic database in which AhR activity was altered. A set of 63 genes was identified (the AhR gene expression biomarker) that was dependent on AhR for regulation after exposure to TCDD or benzo[a]pyrene and includes the known AhR targets *Cyp1a1* and *Cyp1b1*. A fold-change rank-based test (Running Fisher's test; p -value 10^{-4}) was used to evaluate the similarity between the AhR biomarker and a test set of 37 and 41 biosets positive or negative, respectively for AhR activation. The test resulted in a balanced accuracy of 95%. The rank-based test was used to identify factors that activate or suppress AhR in an annotated mouse liver/mouse primary hepatocyte gene expression database of ~1850 comparisons. In addition to the expected activation

*Corresponding author at: Integrated Systems Toxicology Division, National Health and Environmental Effects Research Lab, US Environmental Protection Agency, 109 T.W. Alexander Dr. MD-B143-06 Research Triangle Park, NC 27711, United States. Tel.: +1 919 541 0092; fax: +1 919 541 0694., corton.chris@epa.gov (J.C. Corton).

Author contributions

KO, NV, and JC analyzed the microarray data.

KO, NV generated and analyzed the RT-PCR data.

RT and DA performed animal experiments and contributed microarray data.

LA, CK and FG performed animal experiments.

KO and JC conceived of the study, participated in the study design and helped to draft the manuscript.

All authors read and approved the final manuscript.

¹Present address: Toray Industries, Inc., Kanagawa, Japan.

²Present address: National Center for Computational Toxicology, US-EPA, Research Triangle Park, NC 27711, United States.

Conflict of interest

None.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.tox.2015.07.005>.

of AhR by TCDD and DLC, AhR was activated by AP20189 and phenformin. AhR was suppressed by phenobarbital and 1,4-Bis[2-(3,5-dichloropyridyloxy)] benzene (TCPOBOP) in a constitutive activated receptor (CAR)-dependent manner and pregnenolone-16 α -carbonitrile in a pregnane $\times$ receptor (PXR)-dependent manner. Inactivation of individual genes in nullizygous models led to AhR activation (*Pxr*, *Ghrhr*, *Taf10*) or suppression (*Ahr*, *Ilt6st*, *Hnf1a*). This study describes a novel screening strategy for identifying factors in mouse liver that perturb AhR in a gene expression compendium.

Keywords

Aryl hydrocarbon receptor; Peroxisome proliferator-activated receptor; Transcript profiling; Liver cancer; Constitutive activated receptor; Keap1; Nrf2; Pregnane X receptor

1. Introduction

Adverse outcome pathways (AOPs) are defined as a series of mechanistically-linked key events starting with a molecular initiating event (MIE) in which a chemical or other stressor interacts with a target proceeding to an adverse outcome in a tissue (Ankley et al., 2010; Pery et al., 2013). A group of AOPs that converge on adverse effects in the liver involve a group of ligand-activated transcription factors including the aryl hydrocarbon receptor (AhR). AhR is a member of the basic region helix-loop-helix/period-aryl hydrocarbon nuclear translocator-simple-minded (bHLH/PAS) family of transcription factors that controls a variety of developmental and physiological events including the induction of drug and xenobiotic metabolizing enzymes, response to hypoxia, and hormone receptor function (Bock and Kohle, 2009; McMillan and Bradfield, 2007). For humans, rats and mice the mechanism for AhR activation has similar features. In its nonligand-bound form, AhR is primarily found in the cytoplasm complexed with a number of proteins including the 90-kDa heat shock protein (Hsp90) (Heid et al., 2000). Upon ligand binding, the AhR complex undergoes a conformational change resulting in complex dissociation, translocation to the nucleus, and subsequent AhR heterodimerization with the aryl hydrocarbon nuclear translocator (ARNT), another member of the bHLH/PAS family (Bock and Kohle, 2009; McMillan and Bradfield, 2007). The AhR/ARNT complex, in cooperation with co-regulator proteins, interact with enhancer sequences in the promoters of target genes called dioxin-responsive elements (DREs), and either stimulate or suppress target gene transcription (Hankinson, 1995; Schmidt and Bradfield, 1996). Several genes involved in phase I (e.g., *CYP1A1*, *CYP1A2*, *CYP1B1*) and phase II (e.g., *NQO1* and *ALDH4*) xenobiotic metabolism contain functional DRE sequences in their promoter regions and are transcriptionally activated by a variety of AhR agonists (Schmidt and Bradfield, 1996).

AhR is activated by 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD; also generically referred to as dioxin), one of the most potent known AhR agonists, as well as dioxin-like chemicals (DLCs) that include other polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and polychlorinated biphenyls (PCBs). Compounds which activate AhR include many other industrial compounds and byproducts, widely used pharmaceuticals, and several classes of chemoprotective phytochemicals such as the

flavonoids, indole-3-carbinol, and related compounds (reviewed in Denison et al., 2011). In general, the ligands share similar structural features, i.e., small, aromatic, and planar. However, the ligands can differ markedly in how they affect AhR function (Safe et al., 2013).

Initial interest in the mechanistic basis for AhR action arose from the relationship between activation of this receptor, carcinogenicity in laboratory animals, and concerns over occupational and environmental exposures to TCDD and DLCs (Knerr and Schrenk, 2006). Chronic exposure to TCDD leads to increased incidence of hepatocellular adenoma, cholangiocarcinoma and cholangioma, predominantly in female rats. In mice, both males and females develop liver tumors following TCDD treatment (Budinsky et al., 2013). Lifetime cancer bioassays in rodents have been carried out for other dioxins, furans, and dioxin-like PCBs including 2,3,4,7,8-pentachlorodibenzofuran, a mixture of hexachlorodibenzo-p-dioxins, and PCB126 and PCB118. These studies are notable for a close relationship between liver histopathology, tumor development and the unusual extent and presentation of the hepatic lesions (Goodman and Sauer, 1992; Hailey et al., 2005). The AOP from chronic TCDD or DLC exposure to liver cancer includes chronic activation of AhR, alteration of genes that control cell fate, inhibition of intrafocal apoptosis, increased cell proliferation of previously initiated foci and proliferation of oval cells (Budinsky et al., 2013). Although not considered hepatocarcinogenic themselves, endogenous ligands of AhR (e.g., kynurenine, kynurenic acid and indoles) may modulate the activation of AhR and thus the hepatocarcinogenic process (Budinski et al., 2013). Short-term liver toxicity is not observed in either *Ahr*-null mice (Gonzalez and Fernandez-Salguero, 1998) or TCDD-resistant Han Wistar/Kupio rats at doses of TCDD that are hepatotoxic in sensitive strains such as Sprague-Dawley rats (Sand et al., 2010). Recent studies with an *AHR*-null rat model demonstrated that AhR was required for TCDD to induce liver weights and liver xenobiotic metabolism genes (Harrill et al., 2013). Congenic mice homozygous for either the *Ahr^{b1}* or *Ahr^d* alleles (encoding an AhR with high or low binding affinity for TCDD, respectively) were used to demonstrate that hepatocellular tumor promotion in response to TCDD was dependent on the status of the *Ahr* locus (Kennedy et al., 2014). Overall, these studies support AhR activation as the MIE in an AOP involving TCDD- or DLC-induced liver cancer.

Increasing evidence points to a role for AhR in fatty liver disease. In humans exposed to dioxin, there is an increased incidence of fatty liver (Lee et al., 2006). Transgenic or pharmacological activation of AhR in mice led to spontaneous fatty liver disease (Lu et al., 2011 a,b) by transcriptionally activating the fatty acid translocase *FAT/Cd36* gene and increasing fatty acid uptake (He et al., 2011). Moreover, TCDD increases the expression of intestinal and hepatic lipid transport genes including *Slc27a*, *Fabp*, *Ldlr*, and *Apob* (Angrish et al., 2012). A mouse strain (B6.D2 strain) encoding a low-affinity AhR was resistant to the effects of a high fat western-style diet including increases in body mass, gonadal fat pad mass, liver to body weight ratios, liver damage and steatosis compared to the parent C57BL/6 (B6 strain), which naturally bears the high-affinity AhR (Kerley-Hamilton et al., 2012). Activation of AhR also increased oxidative stress (Lu et al., 2011a; Dalton et al., 2002), a risk factor for nonalcoholic steatohepatitis (NASH), possibly through decreases in superoxide dismutase activity (*Sod2*) controlled by the mitochondrial NAD-dependent

deacetylase sirtuin 3 (*Sirt3*) (He et al., 2013). TCDD treatment led to decreased expression of the hepatic triglyceride hydrolase carboxylesterase 3 (*Ces3*) in an AhR-dependent manner in mouse liver and forced expression of *Ces3* ameliorated steatohepatitis caused by a methionine and choline deficient diet (Matsubara et al., 2012). In summary, an hypothesized AOP leading to steatohepatitis in mice involves (1) activation of AhR (the MIE); (2) increased expression of fatty acid uptake genes (*Cd36*, *Slc27a*, *Fabp*, *Ldlr*, *Apob*) and decreased *Ces3* expression involved in triglyceride degradation; (3) increased accumulation of triglycerides in hepatocytes (“first hit”); (4) increases in oxidative stress through decreases in Sirt3-controlled *Sod2* (“second hit”); and (5) steatohepatitis.

The ability to predict AhR activation or suppression in gene expression profiles would be useful to identify chemicals and other factors that modulate AhR. In the present study, a gene expression biomarker that predicts AhR activation or suppression in mice was characterized by identifying genes consistently altered by AhR activation in wild-type but not AhR-null mice. This biomarker coupled with a rank-based statistical test was found to accurately predict activation of AhR and was used to screen a mouse liver gene expression compendium of ~1850 comparisons to find chemicals and other factors that modulate AhR activity. The predictions could be used as a starting point to link AhR activation and the adverse outcomes of liver cancer and steatosis.

2. Materials and methods

2.1. Strategy for identification of factors that affect AhR

The methods used in this study are outlined in Fig. 1. A screen for AhR modulators required (1) a gene expression biomarker of AhR-dependent genes, and (2) an annotated database of gene expression profiles of statistically filtered genes (also called biosets). The AhR biomarker is a list of genes with associated fold-change values that reflect average differences in expression between control and treated mice and that require AhR for these differences. A commercially available gene expression database provided by NextBio (www.nextbio.com) was used for screening. The NextBio database contains over 123,000 lists of statistically filtered genes from over 18,800 studies carried out in 16 species (as of June 1, 2015). Each list (bioset) is compared to all other biosets in the database using a fold-change rank-based statistical test called the Running Fishers test which allows an assessment of the overlap in regulated genes and whether those genes are regulated in a similar or opposite manner between two biosets. In this study, only biosets from mouse liver, mouse primary hepatocytes, or hepatocyte-derived cell lines were evaluated. Available information about each bioset was extracted from NextBio and used to populate a compendium of information about the experiments. Each bioset was further annotated using information derived from the original GEO submission and/or the original publication. Importantly, each bioset was annotated for the type (e.g., chemical) and name (e.g., phenobarbital) of the factor examined. To assess AhR activation or suppression, the AhR biomarker was uploaded to the Nextbio database and compared to all biosets in the database using the Running Fishers test. Results of the test were exported and used to populate the annotated compendium with *p*-values of each comparison. Test results were used to determine the accuracy of predictions and to characterize those biosets that achieved statistical

significance. Additional analyses were carried out using an independent database of experiments.

2.2. Animal studies

There were a total of 3 animal studies carried out as part of this investigation. The studies of PFNA in wild-type and PPAR α -null mice (Study 1), pregnenolone-16- α -carbonitrile (PCN) in wild-type and Pxr-null mice (Study 2), and the 12 treatment study in male and female mice have been described previously (Study 3) (Oshida et al., 2015a).

2.3. RNA isolation

Total RNA was isolated and purified from mouse livers according to the mirVana miRNA isolation kit (Ambion, Austin, TX) or by Qiagen RNeasy or AllPrep Mini/Midi kits. The integrity of each RNA sample was determined using an Agilent 2100 Bioanalyzer (Agilent, Foster City, CA), and RNA quantity was determined using a Nanodrop ND-1000 (Thermo Fisher Scientific, Wilmington, DE).

2.4. Microarray analyses

Microarray analyses were performed on the livers of mice as part of this study, and the raw data files and experiment descriptions are archived in GEO. These studies include PFNA in wild-type and PPAR α -null mice (GSE55756), PCN in wild-type and PXR-null mice (GSE55746), and the 12 treatment study (GSE55084). Liver gene expression analysis was performed according to the Affymetrix recommended protocol using Affymetrix Mouse Genome 430 2.0 GeneChips $^{\text{®}}$ containing probes for over 30,000 well-characterized genes. There were 4 or 5 biological replicates from each group that were analysed in each study. Total RNA (5 μ g per sample) was labeled using the Affymetrix $^{\text{®}}$ One-Cycle cDNA Synthesis protocol and hybridized to arrays as described by the manufacturer (Affymetrix $^{\text{®}}$, Santa Clara, CA). Microarray hybridizations were conducted overnight at 45 C while rotating in an Affymetrix hybridization oven. After 16 h of hybridization, the cocktail was removed and the arrays were washed and stained in an Affymetrix GeneChip $^{\text{®}}$ fluidics station 450 according to the Affymetrix-recommended protocol. Arrays were scanned on an Affymetrix GeneChip $^{\text{®}}$ scanner.

2.5. Identification of differentially expressed genes in microarray datasets

All of the Affymetrix cel files from the above studies were first analyzed by Bioconductor SimpleAffy (Wilson and Miller, 2005) to assess sample quality. Cel files from individual studies were normalized using Rosetta Resolver $^{\text{®}}$ version 7.1 Affymetrix Rosetta-Intensity Profile Builder software (Rosetta Inpharmatics, Kirkland, WA). The cel files from each study were normalized as a group. Statistically significant genes were identified by one-way ANOVA with a false discovery rate (Benjamini-Hochberg test) of 0.01. In addition to the studies described here, a number of publicly available studies were analyzed using the same methods to derive statistically significant gene lists or biosets (i.e., gene expression differences between two conditions) (to give a total of ~890). Publicly available sources included Gene Expression Omnibus (GEO) and ArrayExpress. Tables were built using the common probe sets shared by 430A or 430_2 chip types.

2.6. Classification analysis using machine learning methods

Analyses were performed using BRB-ArrayTools version 4.2.1 Stable Release developed by Dr. Richard Simon and BRB-ArrayTools Development Team (<http://linus.nci.nih.gov/BRB-ArrayTools.html>) (Simon et al., 2007). All samples used in the training and testing sets were first log2 normalized using RMA in the RMAExpress software environment (<http://rmaexpress.bmbolstad.com/>). The raw files from the three Affymetrix array types (mouse 430A, mouse 430_2 and mouse 430PM arrays) were normalized separately. Normalized expression values of common probesets (22,626) were then combined into one master file. Prior to classification, probesets were excluded under any of the following conditions: (1) minimum fold change -less than 20% of the expression data values have at least a 1.5-fold change in either direction from the genes median value, (2) variance is in the bottom 75th percentile, or (3) percent missing exceeds 50%. Filtering using these criteria resulted in 5644 probesets used in the classification study. The 7 models used for class prediction included Compound Covariate Predictor, Bayesian Compound Covariate, Diagonal Linear Discriminant Analysis, 1- and 3-Nearest Neighbor Classifications, Nearest Centroid, and Support Vector Machines with Linear Kernel. The models incorporated genes that were differentially expressed at $p = 0.001$ significance level as assessed by the random variance t -test. The prediction error of each model was estimated using 10-fold cross-validation. Two training sets were used for predicting AhR activation. The first training set consisted of samples from wild-type and AhR-null mice and included 12 positives and 23 negatives. The second training set consisted of the same set but lacked the control and treated *Ahr*-null samples and included 12 positives and 18 negatives. The derived classifiers of 110 and 19 probesets were then used to predict AhR activation of the remaining samples. A test set of samples known to be positive or negative, respectively for AhR activation came from a number of studies in which mice or mouse primary hepatocytes were exposed to AhR activators or control substances.

2.7. Identification of an AhR-dependent gene biomarker

Methods for biomarker development are outlined in Fig. 1, left. The usefulness of a number of automated methods to find gene sets predictive of AhR activation was evaluated using BRB-Array Tools (discussed above). However, these methods could not control parameters that would identify the most relevant genes including the inability to filter for absolute AhR-dependence. Many of the genes identified using the automated methods had only partial dependence on AhR (data not shown). Manual methods were used for identification of genes consistently altered by AhR activators and that were AhR-dependent. A list of probe sets that comprise the AhR biomarker was derived using livers of wild-type and *Ahr*-null mice treated with TCDD for 19 h (Tijet et al., 2006), and primary hepatocytes from wild-type and *Ahr*-null mice treated with TCDD or benzo[a]pyrene (B[a]P) for 8 h (Sartor et al., 2009). The strategy was similar to that reported earlier to find genes regulated by SREBP1 (Horton et al., 2003) although with stricter criteria. There were 4–5 biological replicates in each group in the Tijet et al. study and 3 biological replicates in each group in the Sartor et al. study. Three statistical tests were used to generate each of the three lists of AhR-dependent genes. Gene lists were derived from chemical treated wild-type tissues (liver or hepatocytes) compared to wild-type controls and chemical treated *Ahr*-null mice tissues compared to AhR-null controls. Genes were identified which exhibited differences in expression in wild-

type mice but no statistically significant expression in the same direction in the *Ahr*-null mice. This list of genes was then examined for statistical differences between the treated wild-type mice and the treated *Ahr*-null mice. The lists of genes were compared and probe sets were selected based on the following criteria: (1) probe sets exhibited the same direction of change, (2) probe sets that encoded the same gene had identical direction of change after exposure, (3) probe sets were altered in 2 or 3 out of the 3 comparisons, (4) the average fold change for each probe set was $|\pm 1.5|$ -fold, and (5) the probe sets were not also altered in the same direction in gene expression biomarkers for PPAR α , CAR (Oshida et al., 2015a,b), Nrf2, and STAT5b (Oshida et al., in preparation). These criteria for selection of the probe sets insured that the probe sets met the following characteristics: absolute dependence on AhR, robust response, and consistent chemical-independent regulation. The final list of probesets in the AhR biomarker is found in Supplemental File 2.

2.8. Functional analysis of the AhR biomarker genes

The Comparative Toxicogenomics Database (CTD; <http://ctdbase.org/>) was used to find annotated relationships between chemicals and genes. Only the annotations “decreases expression”, or “increases expression” were used. A table was built comprised of genes in the AhR biomarker and annotations from CTD using common gene abbreviations. The full list of genes in the AhR biomarker was analyzed using the Ingenuity Pathways Analysis (IPA) canonical pathway and upstream analysis functions. The significance for canonical pathways was calculated using a right-tailed Fisher’s Exact test by IPA. Upstream analysis function of potential regulators of the AhR biomarker genes was based on the number of differentially expressed downstream genes in the biomarker and quantified using a z-score and *p*-value. All results were exported as excel files and filtered based on *p*-value.

2.9. Assembly of an annotated mouse liver gene expression compendium

The compendium consists of information about each bioset and predictions of AhR activation or suppression. All of the biosets examining gene expression in mouse liver, mouse primary hepatocytes or mouse liver cell lines were annotated for study characteristics allowing systematic assessment of the effect of different factors on AhR. The list of descriptors provided for each of the biosets included study ID (i.e., source), bioset name, classification of perturbant (e.g., chemical, diet, genotype, etc.), chemical or treatment, CAS #, sex, source, array type, and type of genetic model used (e.g., wild-type, knockout/knockdown or transgenic). Biosets from NextBio derived from the same experiments as those described above were removed to eliminate redundancy. Detailed descriptions of each experiment are available at GEO (<http://www.ncbi.nlm.nih.gov/geo/>) or ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>) if archived there or at the referenced publications.

2.10. Evaluation of AhR activation using the Running Fisher’s Test

A rank-based nonparametric analysis strategy called the Running Fisher’s test within the NextBio (<http://www.nextbio.com/>) environment was used for evaluating AhR modulation. The Running Fisher’s test is analogous to the Gene Set Enrichment Analysis (GSEA) method (Lamb et al., 2006; Subramanian et al., 2005). The Running Fisher’s algorithm differs from GSEA in the assessment of the statistical significance, as *p*-values are computed by a Fisher’s exact test rather than by permutations (Kupersmidt et al., 2010 for details).

This normalized ranking approach enables comparability across data from different studies, platforms, and analysis methods by removing dependence on absolute values of fold-change, and minimizing some of the effects of normalization methods used, while accounting for the level of genomic coverage by the different platforms. The AhR biomarker was imported into NextBio and then compared to all biosets using the Running Fisher's test. Results of the Running Fisher's test (i.e., comparisons of the AhR biomarker to each bioset in NextBio) were exported and used to populate the database with p -value, and correlation direction.

2.11. Classification prediction of AhR activation using the Running Fisher's Test

Biosets from microarray experiments in which the AhR activation state was known (either on or off) were identified from studies that examined the transcriptional effects of AhR activators as well as other studies in which AhR was known not be altered (E-MEXP-1231, E-MEXP-2209, E-MEXP-2636, E-TABM-1139, GSE10082, GSE10769, GSE11796, GSE15859, GSE17925, GSE24751, GSE33248, GSE35058, GSE44088, GSE55084). The final number of biosets evaluated was 45 positives and 25 negatives. Unlike the 7 machine learning classification methods described above, optimal conditions for classification were not derived from gene behavior as the biomarker was fixed. A p -value 10^{-4} was selected as our cutoff for significance. In preliminary studies, our group constructed biomarkers for three transcription factors (i.e., AhR, CAR, PPAR α). The p -values from the Running Fisher's test for each of the three biomarkers across the ~1850 biosets were subjected to a Benjamini-Hochberg multiple test correction of $\alpha = 0.001$ resulting in a corrected p -value = $9.4\text{E-}5$ that was rounded to a p -value = $1\text{E-}4$. Classification tests were subsequently performed using this p -value cutoff. This cutoff allowed accurate prediction of CAR and PPAR α activation of 97 and 98%, respectively (Oshida et al., 2015a,b). The ability of the biomarker to predict AhR activation was compared to that using only *Cyp1a1* and *Cyp1a2* gene expression. AhR was called activated if both *Cyp1a1* and *Cyp1a2* genes were increased in expression (fold change ≥ 1.2 -fold). The same biosets were used in the analysis using the AhR biomarker and the analysis using the expression of the genes.

Additionally, the mouse AhR biomarker was evaluated for ability to identify AhR activators in a rat compendia. Chemicals that were true positives were those previously identified in the study of Hu et al. (2007) which overlapped with those in a rat liver compendium built for assessment of transcription factor analysis in rat liver or rat primary hepatocytes. To be called AhR activators, chemicals had to exhibit activity in three assays as part of the Hu et al. study including binding to a DRE using electromobility shift assays, AhR binding, and increased expression of *CYP1A1* gene expression. There were 195 biosets that were from the livers of rats or rat primary hepatocytes exposed to one of 11 of these chemicals. The cutoff was the same as above (p -value $1\text{E-}4$). In a similar analysis, the mouse AhR biomarker was evaluated for ability to identify AhR activators in a human compendia. There were 39 biosets from human primary hepatocytes exposed to one of three AhR activators: benzo[a]pyrene, hexachlorobenzene, TCDD. We used the same cutoff as above.

2.12. Additional computational analyses

Principal components analysis (PCA) was evaluated using unfiltered fold-change values of the AhR biomarker probe sets from biosets examined by Affymetrix 430_2 or 430A chips

using Eisen Lab Cluster (<http://rana.lbl.gov/EisenSoftware.htm>) and visualized by SigmaPlot. Heat maps were generated using Eisen Treeview software (<http://rana.lbl.gov/EisenSoftware.htm>).

2.13. Studies used for Real-Time RT-PCR

Livers used for the real-time reverse transcription-PCR (RT-PCR) experiments were from the following studies: WY (50mg/kg/day for 3 days) and AGN (3 mg/kg/day for 3 days) in wild-type and *Ppara*-null mice (Anderson et al., 2004), phenobarbital (0.085%, w/w diet for 28 days) in wild-type and *Car*-null mice (Ross et al., 2009), TCDD (one 100 µg/kg injection and sacrifice 2 weeks later) in wild-type and *Ahr*-null mice (Matsubara et al., 2012), and oltipraz (75 mg/kg/day for 4 days) in wild-type and *Nrf2*-null mice (Aleksunes and Klaassen, 2012). Study 1 and 2 (above) tissues were also used.

2.14. Evaluation of selected genes by real-time RT-PCR

The levels of expression of selected genes were quantified using RT-PCR analysis. Briefly, total RNA was reverse transcribed with murine leukemia virus reverse transcriptase and oligo(dT) primers. The forward and reverse primers for selected genes (available upon request) were designed using Primer Express software, version 2.0 (Applied Biosystems, Foster City, CA). The SYBR green DNA PCR kit (Applied Biosystems, Foster City, CA) was used for real-time PCR analysis. The relative differences in expression between groups were expressed using cycle threshold (Ct) values as follows. The Ct values of the genes were first normalized with β -actin and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) of the same sample. Assuming that the Ct value is reflective of the initial starting copy and that there is 100% efficiency, a difference of one cycle is equivalent to a two-fold difference in starting copy. Means and SE (n = 4) for RT-PCR data were calculated by Student's *t* test. The level of significance was set at $p < 0.05$.

3. Results

3.1. Classification analysis using machine learning algorithms

AhR activation was predicted using 7 classification models as detailed in the Methods. To determine the effect of *Ahr*-null samples on prediction, two training sets were used in the prediction models including the samples from livers of wild-type and *Ahr*-null mice treated with TCDD or B[a]P (Tijet et al., 2006; Sartor et al., 2009) and the same dataset lacking the control and treated *Ahr*-null samples. The derived classifiers of 110 and 19 probesets, respectively were then used to predict AhR activation of test samples. A test set of 21 and 219 samples positive or negative, respectively for AhR activation came from studies in which mice or mouse primary hepatocytes were exposed to known AhR activators or vehicle controls. The models using the wild-type and *Ahr*-null samples in the training set had generally poor sensitivity (sensitivity range, 52–71%; mean, 68%) but excellent specificity (specificity range, 85–100%; mean, 95%) (Supplemental File 1). The models using only the wild-type samples as the training set had lower sensitivity (range, 52–62%; mean, 58%) but about the same specificity (range, 92–100%; mean, 96%), indicating wild-type vs. null comparisons of chemical treatments have some value to improve sensitivity in classification

predictions. Because of the low sensitivity, none of the models were thought to be adequate for predicting AhR activation of additional samples.

3.2. Construction and analysis of an AhR biomarker

AhR biomarker genes were identified as described in the Methods using the biosets mentioned above (livers of wild-type and *Ahr*-null mice treated with TCDD for 19 h (Tijet et al., 2006) or primary hepatocytes from wild-type and *Ahr*-null mice treated with TCDD or B[a]P for 8 h (Sartor et al., 2009)). A summary of the methods used to create the biomarker are shown in Fig. 1, left. A total of 90 probe sets (78 with increased expression and 12 with decreased expression, collapsing to 63 genes) were identified which exhibited similar regulation by two or three out of the three comparisons. The full list of genes is found in Supplemental File 2. The expression of the genes in the *Ahr* biomarker was examined in treated wild-type and AhR-null mice for the three comparisons. Fig. 2A shows the complete dependence of chemical-induced expression on AhR. Many of the genes are well known targets of AhR including *Cyp1a1*, *Cyp1b1* and *Aldh3a1* (reviewed in Dietrich and Kaina, 2010).

To determine whether the genes in the AhR biomarker were previously identified as regulated by AhR activators, the Comparative Toxicogenomics Database (CTD; <http://ctdbase.org/>) was used to find published relationships between chemicals and the biomarker genes in mice. Fig. 2B shows the gene-chemical interactions for 4 known AhR activators comparing effects on the 58 genes annotated in CTD. Data in the figure was derived from studies in which the chemical was given to mice and the indicated gene was shown to be altered. Yellow indicates that one or more studies found the gene to be increased in expression, and blue indicates that one or more studies found the gene to be decreased in expression. Most of the gene-chemical interactions for the genes in the biomarker have been annotated for TCDD and B[a]P, which are commonly used reference chemicals for AhR activation. Far fewer interactions have been annotated for methylcholanthrene and β -naphthoflavone. In general, the direction of expression change caused by exposure to the chemicals assessed from the literature annotations was consistent with the direction of change of the genes in the biomarker. There were a number of examples of two or more citations in which a gene exhibited increased and decreased regulation (green) or regulation opposite to that of the biomarker. These inconsistencies in expression changes with the biomarker may be attributed to effects observed in tissues other than the liver as CTD does not document the tissue origin of the change. Overall, the majority of the AhR biomarker genes exhibited directional changes in chemical-induced expression consistent with annotated findings in the literature.

The AhR biomarker genes were evaluated for canonical pathway enrichment by Ingenuity Pathway Analysis (IPA) (Fig. 2C). The pathway “Aryl Hydrocarbon Receptor Signaling” was enriched as expected. Top enriched pathways included degradation of nicotine, bupropion, acetone, ethanol, melatonin as well as estrogen biosynthesis. The upstream analysis function of IPA predicted that a number of transcription factors regulate the biomarker genes (Fig. 2D). The top scoring transcription factor was AhR (p -value = 5.15×10^{-7}). Other transcription factors included ARNT, GR (NR3C1), and ER α (ESR1).

Behavior of AhR biomarker genes *Cyp1a1* and *Cyp1b1* was evaluated in the livers of mice exposed to activators of AhR, CAR, Nrf2, PPAR α , PXR and RXR. The results are presented in Supplemental File 1. Overall, comparison of gene expression in a battery of wild-type and nullizygous mice shows the specificity of the *Cyp1a1* and *Cyp1a2* gene expression consistent with regulation by AhR but not other xenobiotic-activated transcription factors.

3.3. A rank-based strategy to predict AhR activation

In addition to the machine learning algorithm strategies discussed above, other approaches were considered to predict AhR activation in genomic databases. A rank-based enrichment analysis strategy called the Running Fisher's test within the NextBio environment (Kupersmidt et al., 2010) was evaluated for prediction. Using the Running Fisher's test, the AhR biomarker was compared to gene lists resulting in a correlation direction (positive or negative) and an associated *p*-value of the similarity (Kupersmidt et al., 2010). Because the Running Fisher's test uses similarity as a metric, it could be assumed that biosets which have similarity to each other (based on a low *p*-value) would appear similar. To visualize the relationship between the Running Fisher's test *p*-value and the expression of genes in the biomarker, 371 biosets of statistically filtered genes were evaluated for similarity to the AhR biomarker and then sorted by *p*-value. The left of Fig. 3A shows that for biosets which had a positive correlation to the biomarker, the lower the *p*-value the more the bioset exhibits similarities to the biomarker due to direction and relative magnitude of the changes. These biosets included studies in which B[a]P, DMBA or TCDD were administered to wild-type mice or to mouse primary hepatocytes and which were expected to exhibit similar behavior. The right of the figure shows a smaller group of biosets which exhibited negative correlation to the biomarker. The biosets on the far right exhibited the lowest *p*-values for negative correlation. These biosets included comparisons between AhR-null mice vs. wild-type mice. In general, these biosets exhibited a pattern of gene expression that was opposite to that of the biomarker. Given that AhR can be suppressed by a number of factors (see below), the biosets with negative correlation to the AhR biomarker genes may be hypothesized to reflect AhR suppression.

To determine the predictive behavior of the biomarker, biosets from mice treated with AhR activators were examined. As expected, the biosets from the wild-type comparisons used to build the biomarker exhibited statistically significant similarity (*p*-values 10^{-18}), whereas the biosets from the corresponding *Ahr*-null comparisons were not significant (Fig. 3B). The behavior of the biomarker was further examined by assessing AhR activation in three strains of mice given TCDD at 3 dose levels for 24 h (from E-MEXP-1231). Fig. 3C (left panel) shows consistent activation of AhR at the highest dose level in all strains with decreasing significance in the correlation at the two lower doses. The ability of the biomarker to identify AhR activation in *in vitro* cultures was also assessed. Primary mouse hepatocytes were dosed with 1 nM TCDD and transcriptional effects were measured between 12 and 48 h (from E-MEXP-2209/2539). AhR exhibited activation at all time points (Fig. 3C, middle panel). Responses to TCDD were compared between hepatocytes from wild-type mice and those from AhR-null mice expressing a human AhR (hAhR; from GSE17925). AhR activation was observed in both genotypes (Fig. 3C, right panel). Lastly, mouse primary hepatocytes dosed with 5–60 μ M of B[a]P for 24 h (E-TABM-1139) exhibited significant

AhR activation at 15 μ M and above (Fig. 3D). Thus, the biomarker correctly identified activation of AhR by TCDD and B[a]P under a variety of *in vivo* and *in vitro* conditions.

The ability of the biomarker to predict AhR activation was examined in an independent study to determine whether the biomarker could differentiate compounds that are known activators of AhR from those that activate other transcription factors. Male and female mice were administered 12 different chemicals or biological agents. These included those that principally activate AhR (TCDD, β -naphthoflavone), but also included those that activate CAR (aroclor-1260, PCB-153, phenobarbital), or PPAR α (ciprofibrate, WY). As expected, TCDD in male mice and TCDD and β -naphthoflavone in female mice activated AhR (Fig. 3E). The activation of AhR in male mice by β -naphthoflavone did not reach significance and could be due to the relatively modest transcriptional response observed after exposure (data not shown). The other reference chemicals did not activate AhR, as expected. Interestingly, phenylhydrazine caused suppression of AhR in female but not male mice. Overall, the AhR biomarker correctly identified compounds which activate AhR. No known sex differences in toxicity to β -naphthoflavone in mice have been published. However, sex differences to the acute effects of TCDD toxicity in mice have been reported with females being more resistant (Pohjanvirta et al., 2012).

Biosets of known positives and negatives for AhR activation were examined by principal components analysis (PCA) to determine if the expression behavior of the genes in the AhR biomarker can separate the two groups. PCA was evaluated using unfiltered fold-change values of the AhR biomarker probe sets from experiments examined by Affymetrix 430.2 or 430A chips. Fig. 3F shows the clear separation between the known positives (blue, purple and red dots within the red oval) and the known negatives (black and green dots within the green oval).

Using a larger number of biosets, a classification analysis was performed (Fig. 3G). Criteria for classification included a p -value < 0.0001 . The final number of biosets evaluated was 45 positives and 25 negatives. The biomarker had 89% sensitivity and 100% specificity, giving a balanced accuracy of 95%. The 10 false negatives were from four studies. Three of the false negatives came from a study in which Muta mice were given B[a]P (GSE24907). As these mice were sacrificed 3 days after the last dose, it is possible that B[a]P effects were no longer apparent after 3 days of no treatment. Other explanations for negative calls could include transcriptional response differences between wild-type and Xpa/p53-null strains (E-TABM-1139). Despite these false negatives, the biomarker resulted in an excellent balanced accuracy. Preliminary analyses (Supplemental File 1) showed that the AhR biomarker was a better predictor of AhR activation than the expression of two AhR target genes, *Cyp1a1* and *Cyp1a2*. Using the expression of the two genes, the test resulted in a balanced accuracy of 89% with a specificity of 100% but a sensitivity lower than that of the AhR biomarker (79%). The biomarker was not useful to predict AhR activation in rat livers, rat primary hepatocytes, or human primary hepatocytes (Supplemental File 1).

3.4. Analysis of a mouse liver gene expression compendium

A mouse liver gene expression compendium was constructed as detailed in the Methods. The compendium consisted of gene expression changes in the livers of mice, primary

hepatocytes, or hepatocyte-derived cell lines by diverse factors. The compendium contains ~1850 biosets of gene expression changes between control and experimental states including ~470 chemical, ~450 gene, ~220 diet, ~100 hormone or cytokine, ~90 life stage, ~90 stress and ~120 strain comparisons.

Using the running Fishers test, a total of 89 biosets were classified as affecting AhR including 64 activating and 25 suppressing AhR (p -value 0.0001) (Fig. 4). The distribution of the biosets indicates that out of all of the factors examined, chemicals had the greatest effect on AhR.

3.5. Chemical activation of AhR

The distribution of $-\log(p\text{-values})$ across 503 chemical comparisons representing exposure to ~150 chemicals is shown in Fig. 5A and shows that 46 comparisons (17 chemicals) significantly activated AhR and 10 comparisons (7 chemicals) significantly suppressed AhR. In addition to the chemicals used to create the biomarker (TCDD and B[a]P), a number of chemicals were found to significantly activate AhR. The chemicals include the known activators 7,12-dimethylbenzanthracene (DMBA) (Galvan et al., 2005) and aroclor 1254 (Madra and Smith, 1992; Schaap et al., 2012) (Fig. 5B). In our study, male BALB/C mice fed a diet supplemented with 0.5% quercetin showed activation of AhR (from GSE38136). Quercetin was found to be an indirect activator of AhR, possibly by inhibiting the metabolic turnover of the endogenous AhR ligand 6-formylindolo[3,2-b]carbazole (FICZ) (Mohammadi-Bardbori et al., 2012).

A number of novel chemicals were identified as activators of AhR.

- AP20187 activated AhR. AP20187 was used to increase phosphorylation of the translation initiation factor eIF2 α which led to activation of ATF4, lipogenesis and steatosis (Oyadomari et al., 2008). There is evidence that AhR activates ATF4 in human HepG2 cells (Terashima et al., 2013). Thus, it is possible that there is an overlap in the transcript profiles after AhR activation or after activation of a downstream target, namely ATF4.
- Acetaminophen (APAP) exposure for 3 h led to AhR activation in DBA/2J mice but not in three other strains of mice used in the study (GSE17649; Liu et al., 2010). Compared to these other strains, the DBA/2J mice exhibited the greatest induction of cytotoxicity as assessed using the marker ALT but were resistant to becoming moribund (Liu et al., 2010). The effect of APAP in this strain was transient as AhR activation was not observed after 6h of treatment (data not shown). Given that AhR can be activated under conditions of low glucose (Terashima et al., 2013) and that plasma glucose in mice was significantly reduced after APAP administration (Aubert et al., 2012), it is possible that under the conditions of the Liu et al. (2010) study, glucose levels were suppressed leading to AhR activation. How low glucose levels lead to AhR activation is not known.
- Phenacetin was found to activate AhR in mouse primary hepatocytes after exposure to a high dose (1.5 M) for 48 h (E-MEXP-2636). We speculate that

phenacetin which is metabolized to APAP activates AhR similar to the mechanism of APAP as discussed above.

- Lipopolysaccharide (LPS) administered to mice for 3h activated AhR (dataset from Shaw et al., 2009). There is evidence that LPS exposure can increase the expression of AhR and ARNT in mouse splenocytes and splenic B cells (Marcus et al., 1998), but any increases in the expression of AhR or ARNT, if they occur in the mouse liver would likely not occur until after 3 h of exposure. We speculate that under the conditions of AhR activation glucose levels were suppressed. In fact, LPS exposure decreased blood glucose within 2 h after LPS injection (Raetzsch et al., 2009), making it plausible that AhR is transiently activated through decreases in glucose.
- The FXR activator GW4064 activated AhR in the livers of liver-specific FXR-null mice after a three day exposure (GSE54557). There is no evidence in the literature which indicates that GW4064 directly activates AhR. An examination of the genes altered by GW4064 in liver-specific FXR-null mice showed a significant overlap in the similarity of the gene expression pattern (p -value = $4.3E-21$) with the same bioset described above (mice treated with AP20187), indicating that the mechanism underlying the AhR activation under the two exposure conditions is similar.
- Exposure to 40 mg/kg of 70 nm silica dioxide (SiO₂) particles for 6 hr led to AhR activation (from GSE30861; Lu et al., 2011 b). Silica dioxide nanoparticles when given intravenously mainly localize to the liver and result in disturbances in energy, amino acid, lipid, and nucleotide metabolism that are linked to associated hepatotoxicity (Lu et al., 2011b). Other biosets from this experiment exhibited increased AhR activation that approached significance (p -values <0.001) after exposure to 10 or 20 mg/kg of the 70 nm particles and to 10 mg/kg of the 30 nm particles. Given that these particles, like LPS and APAP discussed above caused cytotoxicity, we speculate that glucose levels may be disrupted leading to activation of AhR.
- The PPAR β activator GW501516 and the PPAR γ activator rosiglitazone both at 10 mg/kg caused AhR activation at 8 d but not 4 d or 15 d (from GSE27948). In addition, WY activated AhR in PPAR α -null mice after 6h (GSE8396). WY can also activate the other PPAR subtypes especially in the absence of PPAR α (Corton and Lapinskas, 2005). A number of endogenous lipid molecules activate AhR including lipoxin A4 (Denison and Nagy, 2003). Exposure to rosiglitazone induced brain 5-lipoxygenase (5-LO) expression in ischemic rat brain. Rosiglitazone also increased cerebral lipoxin A4 levels (Sobrado et al., 2009). Six prostaglandins (prostaglandin B2, D2, F3 α , G2, H1 and H2) were also found to be endogenous inducers of AhR (Denison and Nagy, 2003). PPARs regulated the expression of cyclooxygenase 2 levels in a number of cells and tissues (Meade et al., 1999; Aleshin et al., 2009). Thus, it is possible that these PPAR activators through regulation of either 5-LO or prostaglandin synthesis genes lead to increases in the levels of endogenous activators of AhR.

3.6. Suppression of AhR by CAR and PXR activators

A number of CAR or PXR activators suppressed AhR. TCPOBOP suppressed AhR in wild-type but not CAR-null mice; phenobarbital also suppressed AhR that approached significance in wild-type mice (Fig. 5C, left panel) (from GSE40120). Phenobarbital administered to male but not female mice for 2 or 7 days (from GSE54597) caused significant suppression of AhR at the highest dose (Fig. 5C, right panel). Treatment with the PXR activator PCN in wild-type but not Pxr-null mice resulted in AhR suppression (Fig. 5D, left) (this study). The PXR activator “Compound 13” (C13) exposure in wild-type but not PXR-null mice caused suppression of AhR which approached significance (from GSE23780). Additionally, Pxr-null mice exhibited AhR activation (Fig. 5D, right). Other chemicals suppressed AhR including phenylhydrazine (Fig. 3E, 100 mg/kg i.p. once; p -value = $2.0\text{E}-03$ for male, $3.7\text{E}-06$ for female), myclobutanil (2000 ppm in the feed for 90 d; p -value = $9.8\text{E}-05$), tridimefon (1800ppm in feed for 4d or 90d; p -value = $5.3\text{E}-5$ or $4\text{E}-6$, respectively), and bezafibrate (5000ppm in the feed for 5 d; p -value = $5.1\text{E}-5$). All of these chemicals activate CAR and/or PXR based on studies with CAR and PXR biomarkers (Oshida et al., 2015b; unpublished) or from the original microarray analyses (Ward et al., 2006; Hester et al., 2012). Overall, the results are consistent with CAR and PXR activator exposure at high doses leading to suppression of AhR.

3.7. Effects of nutritional status on AhR activation

The effects of synthetic triglycerides and glucagon on AhR are discussed in Supplemental File 1.

3.8. Genetic activation or suppression of AhR

Effects of specific genetic changes on AhR were examined in 448 comparisons in mouse models encompassing 194 genes including those in which single genes were over-expressed (45 biosets) or inactivated either by homologous recombination or by knockdown technologies (395 biosets). There were 3 biosets that activated and 5 biosets that suppressed AhR (p -value 10^{-4}).

There were 3 biosets from *Ahr*-null to wild-type mouse comparisons, all of which resulted in suppression of AhR (Fig. 6, left panel). This result is consistent with AhR constitutively regulating the biomarker genes, presumably because of the presence of a low level of constitutive activation.

AhR was activated in three nullizygous models (Fig. 6, right) in addition to the *Pxr*-null bioset described above.

- c-Jun is a transcription factor activated by the stress-activated protein kinase/c-Jun N-terminal kinase pathway in response to extracellular signals and cytokines. Expression of a dominant negative form of c-Jun in ApoE-null mice (Drosatos et al., 2007) resulted in AhR activation (p -value = $5\text{E}-5$). C-Jun is a target gene for AhR and induction depends on activation of p38-mitogen-activated protein kinase (MAPK) through AhR (Carsten et al., 2005). B[a]P increased phosphorylation of MEK and p38-MAPK, and increased AP-1 DNA binding and phosphorylation of c-Jun (Oesterling et al., 2008). Although no published studies

could be found which examined direct interactions between c-Jun and AhR proteins, we hypothesize that c-Jun negatively regulates AhR transcriptional activity.

- *Ghrhr* (growth hormone releasing hormone receptor; little) expressed exclusively in the pituitary, encodes a receptor for growth hormone releasing hormone. Activation of the receptor leads to synthesis and release of growth hormone. AhR was activated in the livers of *Ghrhr*-null mice at 3 months of age but not significantly in older *Ghrhr*-null mice. *Ghrhr*-null mice although dwarf in stature are long-lived and exhibit increased expression of xenobiotic metabolism genes which have been hypothesized to contribute to their longevity phenotype (Amador-Noguez et al., 2007).
- *Taf10* (TAF10 RNA polymerase II, TATA box binding protein (TBP)-associated factor) is a component of the basal transcription machinery. Liver-specific *Taf10*-null mice exhibit severe growth failure during the second week after birth, more than 50% reduction in body weight at day P30, and premature death at ~P35 (Kamileri et al., 2012). The livers of these mice revealed a uniform accumulation of triglycerides and glycogen resulting in a fatty liver appearance with unusually large glycogen depots (Kamileri et al., 2012). AhR was activated in *Taf10*-null mice although the activation did not reach significance (p -value = $4E-4$), both *Cyp1a1* and *Cyp1b1* were activated > 6-fold. No evidence for direct interactions between AhR and *Taf10* could be found. However, the human AhR protein was found to interact with components of the general transcription machinery including other TAF proteins TAF4, TAF6, and TAF9 (Watt et al., 2005). Further work is needed to determine if AhR and *Taf10* interact in the mouse liver.

AhR suppression was found in two nullizygous models (Fig. 6, right panel).

- The gp130 interleukin 6 signal transducer (*Il6st*; *Gp130*; *oncostatin M receptor*) encodes a signal transducer shared by many cytokines, including interleukin 6 (IL6), ciliary neurotrophic factor (CNTF), leukemia inhibitory factor (LIF), and oncostatin M (OSM). This protein functions as part of a cytokine receptor complex. *Il6st*-null mice exhibited decreases in AhR activation (p -value = $4E-5$). Mice nullizygous for *Il6st* showed a number of liver abnormalities including an increase in lipid vacuoles (Betz et al., 1998).
- The *Hnf1a*-null mice exhibited decreases in AhR activation (p -value = $5E-5$). The livers of *Hnf1a*-null mice showed progressive liver damage but no reported increases in steatosis (Lee et al., 1998). The reason for the AhR suppression is elusive given the paucity of information about any interactions between AhR and *Hnf1a*. We hypothesize that many of the genes in the AhR biomarker are also regulated by *Hnf1a*. To support this, UDP-glucuronosyltransferase 1a7 (*UGT1A7*) in rat hepatocytes was found to be regulated by both AhR and HNF1a (Metz et al., 2000).

4. Discussion

A biomarker-based approach was used to identify factors in the mouse liver that have an impact on AhR predicted by the expression behavior of AhR-regulated genes. To create an AhR biomarker, genes that were consistently activated or repressed in an AhR-dependent manner were identified using a stringent set of criteria, resulting in a final list of 90 probe sets collapsing to 63 genes (Fig. 2A). The AhR biomarker coupled with a rank-based statistical test reliably predicted AhR activation (Fig. 3). The biomarker was able to distinguish between activation of AhR and other chemically-induced transcription factors (Fig. 3E). Our formal test to predict activation demonstrated a balanced accuracy of 95% (Fig. 3G) exceeding the predictions from a number of machine learning algorithms (Supplemental File 1). The liver biomarker will be a useful tool for predicting AhR effects in future genomic studies in the mouse liver. There are a number of uncertainties in the extent of the application of the biomarker to predict AhR effects. Although there are a minor number of AhR gene targets which exhibit similar regulation across tissues and species (Bock and Kohle, 2009), reliable biomarkers might have to be built using treated and control samples from the tissue and species of interest. We found that the mouse AhR biomarker can identify only a small fraction of the biosets from rat liver, rat hepatocytes or human hepatocytes that have been exposed to well-known AhR activators (Supplemental File 1). Additionally, the number of genes in the biomarker may not lend the biomarker useful in high throughput screens. Future studies might focus on the ability of smaller sets of genes in the biomarker to predict AhR effects. It should be noted that relying on a small number of genes for prediction is potentially problematic as there is greater dependence of predictions on fewer genes. Activation of one of the biomarker genes (*Cyp1a1*) was shown not to strictly correlate with other measures of AhR activation such as reporter gene activation, ligand binding and electrophoretic mobility shift changes in rat liver (Hu et al., 2007). However, using multiple AhR-regulated genes in a test such as in our biomarker may overcome potential deficiencies of individual genes to predict AhR status. Our preliminary examination of the usefulness of *Cyp1a1* and *Cyp1a2* as marker genes for AhR activation showed that expression of these genes approached the balanced accuracy of the AhR biomarker itself (Supplemental File 1).

Comparison of the AhR biomarker to a compendium of liver gene expression biosets uniquely identified diverse factors that when perturbed, resulted in effects on AhR (summarized in Fig. 7A and B). Activation through direct binding to AhR by chemicals is the most widely recognized mechanism by which AhR is activated. The biomarker accurately predicted activation by typical direct activators including TCDD, B[a]P, DMBA, and aroclor 1254 (Figs. 3 and 5B). Other compounds not previously known to be activators of AhR were found in our study (Fig. 5B). Under the different conditions of exposure, many of these compounds are known to alter lipid homeostasis. There is evidence that fatty acids, eicosanoids as well as glucose activate or suppress AhR (Denison et al., 2011; McMillan and Bradfield, 2007; Oesch-Bartlomowicz and Oesch, 2009; Dabir et al., 2008; Terashima et al., 2012), but the mechanisms leading to activation of AhR by these molecules remains elusive. Our studies found that glucagon and synthetic triglycerides composed of C10:0, C18:1, and C18:3 fatty acids activated AhR (Supplemental File 1). Curiously, AhR was only activated by

these triglycerides in PPAR α -null mice but not in wild-type mice. Exposure to the synthetic triglycerides in the absence of a functional PPAR α , would lead to prolonged exposure to the TG or metabolites as the mice are not able to activate fatty acid transport and catabolism genes through PPAR α . Thus, it is possible that under these unique conditions, metabolic intermediates accumulate leading to the activation of AhR.

Chemicals and gene mutations decrease AhR activity (Fig. 7B). Decreases in activity could come through decreases in the expression of the *Ahr* gene itself as seen in AhR-null vs. wild-type mouse comparisons which had a strong negative correlation to the AhR biomarker (Fig. 6). Given that PXR, CAR and AhR regulate an overlapping set of xenobiotic metabolism genes (Moreau et al., 2008), activation or suppression of one transcription factor likely has effects on the others. In our comparisons, evidence for cross-talk between these xenobiotic-regulated factors comes from AhR suppression after exposure to CAR and PXR activators, albeit at high dose levels (Fig. 5C and D). Consistent with these findings, AhR was activated in Pxr-null mice. Taken together, the results indicate that CAR and/or PXR activation leads to suppression of AhR. Suppression could come about through competition for essential cofactors or possibly through the regulation of endogenous activators of AhR.

In summary, these studies provide a useful reference database for chemicals and other factors that can affect AhR. The comprehensive identification of factors that lead to AhR activation or suppression will assist in building cumulative risk models that can incorporate the contribution of chemical-independent factors to predict liver adverse outcomes such as cancer and steatosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was carried out as part of the EPA Chemicals Safety for Sustainability cancer and steatosis AOP projects. We thank Drs. Julian Preston and Charlene McQueen for support, Dr. Jennifer Fostel for archiving the results in CEBS, Dr. William Ward for guidance in analyzing microarray data, Drs. Lyle Burgoon, Susan Hester, Charles Wood and Sheau-Fung Thai for review of the manuscript and Drs. Oliver Hankinson, and Ivan Rusyn for livers from studies carried out in their labs. The information in this document has been funded in part by the U.S. Environmental Protection Agency. It has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Abbreviations:

AhR	aryl hydrocarbon receptor
AOP	adverse outcome pathway; B[a]P, benzo[a]pyrene
CAR	constitutive activated receptor
DLC	dioxin-like compound
DMBA	7,12-dimethylbenzanthracene

DRE	dioxin responsive element
ERα	estrogen receptor α
FXR	farnesoid X receptor
GR	glucocorticoid receptor
MIE	molecular initiating event
Nrf2	nuclear factor (erythroid-derived 2)-like 2
PB	phenobarbital
PCN	pregnenolone-16- α -carbonitrile
PCR	polymerase chain reaction
PFOA	perfluorooctanoic acid
PPAR	peroxisome proliferator-activated receptor
PXR	pregnane X receptor
STAT5b	signal transducer and activator of transcription 5B
TCDD	2,3,7,8-tetrachlorodibenzo-p-dioxin
TCPOBOP	1,4-bis[2-(3,5-dichloropyridyloxy)] benzene
WY	WY-14,643

References

- Aleksunes LM, Klaassen CD, 2012 Coordinated regulation of hepatic phase I and II drug-metabolizing genes and transporters using AhR-, CAR-, PXR- PPAR α -, and Nrf2-null mice. *Drug Metab. Dispos* 40 (7), 1366–1379, 10.1124/dmd.112.045112. [PubMed: 22496397]
- Aleshin S, Grabeklis S, Hanck T, Sergeeva M, Reiser G, 2009 Peroxisome proliferator-activated receptor (PPAR)-gamma positively controls and PPAR α negatively controls cyclooxygenase-2 expression in rat brain astrocytes through a convergence on PPAR β /delta via mutual control of PPAR expression levels. *Mol. Pharmacol* 76 (August (2)), 414–424. [PubMed: 19483106]
- Amador-Noguez D, Dean A, Huang W, Setchell K, Moore D, Darlington G, 2007 Alterations in xenobiotic metabolism in the long-lived Little mice. *Aging Cell* 6 (4), 453–470, <http://dx.doi.Org/10.1111/j.1474-9726.2007.00300.x>. [PubMed: 17521389]
- Anderson SP, Howroyd P, Liu J, Qian X, Bahnemann R, Swanson C, Kwak MK, Kensler TW, Corton JC, 2004 The transcriptional response to a peroxisome proliferator-activated receptor alpha agonist includes increased expression of proteome maintenance genes. *J. Biol. Chem* 279 (50), 52390–52398, 10.1074/jbc.M409347200. [PubMed: 15375163]
- Angrish MM, Mets BD, Jones AD, Zacharewski TR, 2012 Dietary fat is a lipid source in 2,3,7,8-tetrachlorodibenzo-rho-dioxin (TCDD)-elicited hepatic steatosis in C57BL/6 mice. *Toxicol. Sei* 128 (2), 377–386, 10.1093/toxsci/kfs155.
- Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrano JA, Tietge JE, Villeneuve DL, 2010 Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem* 29 (3), 730–741, 10.1002/etc.34. [PubMed: 20821501]

- Aubert J, Begrich K, Delannoy M, Morel I, Pajaud J, Ribault C, Lepage S, McGill MR, Lucas-Clerc C, Turlin B, Robin MA, Jaeschke H, Fromenty B, 2012 Differences in early acetaminophen hepatotoxicity between obese ob/ob and db/db mice. *J. Pharmacol. Exp. Ther* 342 (September (3)), 676–687. [PubMed: 22647274]
- Bock KW, Kohle C, 2009 The mammalian aryl hydrocarbon (Ah) receptor: from mediator of dioxin toxicity toward physiological functions in skin and liver. *Biol. Chem* 390 (12), 1225–1235, 10.1515/bc.2009.138. [PubMed: 19747074]
- Corton JC, Lapinskas PJ, 2005 Peroxisome proliferator-activated receptors: mediators of phthalate ester-induced effects in the male reproductive tract? *Toxicol. Sei* 83 (January (1)), 4–17.
- Dabir P, Marinic TE, Krukovets I, Stenina OI, 2008 Aryl hydrocarbon receptor is activated by glucose and regulates the thrombospondin-1 gene promoter in endothelial cells. *Circ. Res* 102 (12), 1558–1565, 10.1161/circresaha.108.176990. [PubMed: 18515748]
- Dalton TP, Puga A, Shertzer HG, 2002 Induction of cellular oxidative stress by aryl hydrocarbon receptor activation. *Chem. Biol. Interact* 141 (1–2), 77–95. [PubMed: 12213386]
- Denison MS, Nagy SR, 2003 Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu. Rev. Pharmacol. Toxicol* 43, 309–334, Epub 2002 Jan 10. Review. [PubMed: 12540743]
- Denison MS, Soshilov AA, He G, DeGroot DE, Zhao B, 2011 Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol. Sei* 124 (1), 1–22, 10.1093/toxsci/kfr218.
- Dietrich C, Kaina B, 2010 The aryl hydrocarbon receptor (AhR) in the regulation of cell-cell contact and tumor growth. *Carcinogenesis* 31 (8), 1319–1328, 10.1093/carcin/bgq028. [PubMed: 20106901]
- Drosatos K, Sanoudou D, Kypreos KE, Kardassis D, Zannis V.I., 2007 A dominant negative form of the transcription factor c-Jun affects genes that have opposing effects on lipid homeostasis in mice. *J. Biol. Chem* 282 (27), 19556–19564, 10.1074/jbc.M700986200 [PubMed: 17456467]
- Flaveny CA, Murray IA, Perdew GH, 2010 Differential gene regulation by the human and mouse aryl hydrocarbon receptor. *Toxicol. Sei* 114 (April (2)), 217–225, 10.1093/toxsci/kfp308, Epub 2009 Dec 31.
- Galvan N, Teske DE, Zhou G, Moorthy B, MacWilliams PS, Czuprynski CJ, Jefcoate CR, 2005 Induction of CYP1A1 and CYP1B1 in liver and lung by benzo(a)pyrene and 7,12-d imethylbenz(a)anthracene do not affect distribution of polycyclic hydrocarbons to target tissue: role of AhR and CYP1B1 in bone marrow cytotoxicity. *Toxicol. Appl. Pharmacol* 202 (3), 244–257, 10.1016/j.taap.2004.06.026. [PubMed: 15667830]
- Gonzalez FJ, Fernandez-Salguero P, 1998 The aryl hydrocarbon receptor: studies using the AHR-null mice. *Drug Metab. Dispos* 26 (12), 1194–1198. [PubMed: 9860927]
- Goodman DG, Sauer RM, 1992 Hepatotoxicity and carcinogenicity in female Sprague-Dawley rats treated with 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD): a pathology working group réévaluation. *Regul. Toxicol. Pharmacol* 15 (3), 245–252. [PubMed: 1509118]
- Hailey JR, Walker NJ, Sells DM, Brix AE, Jokinen MP, Nyska A, 2005 Classification of proliferative hepatocellular lesions in harlan sprague-dawley rats chronically exposed to dioxin-like compounds. *Toxicol. Pathol* 33 (1), 165–174, 10.1080/01926230590888324. [PubMed: 15805068]
- Hankinson O, 1995 The aryl hydrocarbon receptor complex. *Annu. Rev. Pharmacol. Toxicol* 35, 307–340, 10.1146/annurev.pa.35.040195.001515. [PubMed: 7598497]
- Harrill JA, Hukkanen RR, Lawson M, Martin G, Gilger B, Soldatow V, Lecluyse EL, Budinsky RA, Rowlands JC, Thomas RS, 2013 Knockout of the aryl hydrocarbon receptor results in distinct hepatic and renal phenotypes in rats and mice. *Toxicol. Appl. Pharmacol*, <http://dx.doi.org/10.1016/j.taap.2013.06.024>.
- He J, Hu B, Shi X, Weidert ER, Lu P, Xu M, Huang M, Kelley EE, Xie W, 2013 Activation of the aryl hydrocarbon receptor sensitizes mice to nonalcoholic steatohepatitis by deactivating mitochondrial sirtuin deacetylase Sirt3. *Mol. Cell. Biol* 33 (10), 2047–2055, 10.1128/mcb.01658-12. [PubMed: 23508103]

- He J, Lee JH, Febbraio M, Xie W, 2011. The emerging roles of fatty acid translocase/CD36 and the aryl hydrocarbon receptor in fatty liver disease. *Exp. Biol. Med.* (Maywood) 236 (10), 1116–1121, 10.1258/ebm.2011.011128. [PubMed: 21885479]
- Heid SE, Pollenz RS, Swanson HI, 2000 Role of heat shock protein 90 dissociation in mediating agonist-induced activation of the aryl hydrocarbon receptor. *Mol. Pharmacol* 57(1), 82–92. [PubMed: 10617682]
- Hester S, Moore T, Padgett WT, Murphy L, Wood CE, Nesnow S, 2012 The hepatocarcinogenic conazoles: cyproconazole, epoxiconazole, and propiconazole induce a common set of toxicological and transcriptional responses. *Toxicol. Sci* 127 (May (1)), 54–65, 10.1093/toxsci/kfs086, Epub 2012 Feb 14.
- Horton JD, Shah NA, Warrington JA, Anderson NN, Park SW, Brown MS, Goldstein JL, 2003 Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proc. Natl. Acad. Sci. U.S.A* 100 (21), 12027–12032, 10.1073/pnas.1534923100.
- Hu W, Sorrentino C, Denison MS, Kolaja K, Fielden MR, 2007 Induction of Cyp1a1 is a nonspecific biomarker of aryl hydrocarbon receptor activation: results of large scale screening of pharmaceuticals and toxicants in vivo and in vitro. *Mol. Pharmacol* 71 (6), 1475–1486, 10.1124/mol.106.032748. [PubMed: 17327465]
- Kamileri I, Karakasiloti I, Sideri A, Kostas T, Tatarakis A, Talianidis I, Garinis GA, 2012 Defective transcription initiation causes postnatal growth failure in a mouse model of nucleotide excision repair (NER) progeria. *Proc. Natl. Acad. Sci. U.S.A* 109 (8), 2995–3000, 10.1073/pnas.1114941109.
- Kerley-Hamilton JS, Trask HW, Ridley CJ, Dufour E, Ringelberg CS, Nurinova N, Wong D, Moodie KL, Shipman SL, Moore JH, Kore M, Shworak NW, Tomlinson CR, 2012 Obesity is mediated by differential aryl hydrocarbon receptor signaling in mice fed a Western diet. *Environ. Health Perspect* 120 (9), 1252–1259, 10.1289/ehp.1205003. [PubMed: 22609946]
- Knerr S, Schrenk D, 2006 Carcinogenicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in experimental models. *Mol. Nutr. Food Res* 50 (10), 897–907, 10.1002/mnfr.200600006. [PubMed: 16977593]
- Kupersmidt I, Su QJ., Grewal A, Sundaresh S, Halperin I, Flynn J, Shekar M, Wang H, Park J, Cui W, Wall GD, Wisotzkey R, Alag S, Akhtari S, Ronaghi M, 2010 Ontology-based meta-analysis of global collections of high-throughput public data. *PLoS ONE* 5 (September (9)), 10.1371/journal.pone.0013066, pii:el3066.
- Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, et al., 2006 The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313,1929–1935. [PubMed: 17008526]
- Lee CC, Yao YJ, Chen HL, Guo YL, Su HJ, 2006 Fatty liver and hepatic function for residents with markedly high serum PCDD/Fs levels in Taiwan. *J. Toxicol. Environ. Health A* 69 (5), 367–380, 10.1080/15287390500244972. [PubMed: 16455615]
- Lee YH, Sauer B, Gonzalez FJ, 1998 Laron dwarfism and non-insulin-dependent diabetes mellitus in the Hnf-1 alpha knockout mouse. *Mol. Cell. Biol* 18 (5), 3059–3068. [PubMed: 9566924]
- Lu X, Tian Y, Zhao Q, Jin T, Xiao S, Fan X, 2011b Integrated metabolomics analysis of the size-response relationship of silica nanoparticles-induced toxicity in mice. *Nanotechnology* 22 (February (5)), 055101. [PubMed: 21178262]
- Lu H, Cui W, Klaassen CD, 2011a Nrf2 protects against 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)-induced oxidative injury and steatohepatitis. *Toxicol. Appl. Pharmacol* 256 (2), 122–135, 10.1016/j.taap.2011.07.019. [PubMed: 21846477]
- Madra S, Smith AG, 1992 Induction of cytochrome P450 activities by polychlorinated biphenyls in isolated mouse hepatocytes. Influence of Ah-phenotype and iron. *Biochem. Pharmacol* 44 (3), 455–464. [PubMed: 1510696]
- Marcus RS, Holsapple MP, Kaminski NE, 1998 Lipopolysaccharide activation of murine splenocytes and splenic B cells increased the expression of aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator. *J. Pharmacol. Exp. Ther* 287 (December (3)), 1113–1118. [PubMed: 9864300]

- Mathijs K, Brauers KJ, Jennen DG, Boorsma A, van Herwijnen MH, Gottschalk RW, Kleinjans JC, van Delft JH, 2009 Discrimination for genotoxic and nongenotoxic carcinogens by gene expression profiling in primary mouse hepatocytes improves with exposure time. *Toxicol. Sei* 112 (December (2)), 374–384, 10.1093/toxsci/kfp229, Epub 2009 Sep 21.
- Matsubara T, Tanaka N, Krausz KW, Manna SK, Kang DW, Anderson ER, Luecke H, Patterson AD, Shah YM, Gonzalez FJ, 2012 Metabolomics identifies an inflammatory cascade involved in dioxin- and diet-induced steatohepatitis. *Cell Metab.* 16(5), 634–644, <http://dx.doi.org/10.1016/j.cmet.2012.10.006>. [PubMed: 23140643]
- McMillan BJ, Bradfield CA, 2007 The aryl hydrocarbon receptor sans xenobiotics: endogenous function in genetic model systems. *Mol. Pharmacol* 72 (3), 487–498, 10.1124/mol.107.037259. [PubMed: 17535977]
- Meade EA, McIntyre TM, Zimmerman GA, Prescott SM, 1999 Peroxisome proliferators enhance cyclooxygenase-2 expression in epithelial cells. *J. Biol. Chem* 274 (March (12)), 8328–8334. [PubMed: 10075740]
- Metz RP, Auyeung DJ, Kessler FK, Ritter JK, 2000 Involvement of hepatocyte nuclear factor 1 in the regulation of the UDP-glucuronosyltransferase 1A7 (UGT1A7) gene in rat hepatocytes. *Mol. Pharmacol* 58 (2), 319–327. [PubMed: 10908299]
- Moreau A, Vilarem MJ, Maurel P, Pascussi JM, 2008 Xenoreceptors CAR and PXR activation and consequences on lipid metabolism, glucose homeostasis, and inflammatory response. *Mol. Pharm* 5 (1), 35–41, 10.1021/mp700103m. [PubMed: 18159929]
- Oesch-Bartlomowicz B, Oesch F, 2009 Role of cAMP in mediating AHR signaling. *Biochem. Pharmacol* 77 (4), 627–641, <http://dx.doi.org/10.1016/j.bcp.2008.10.017>. [PubMed: 19013136]
- Oesterling E, Toborek M, Hennig B, 2008 Benzo[a]pyrene induces intercellular adhesion molecule-1 through a caveolae and aryl hydrocarbon receptor mediated pathway. *Toxicol. Appl. Pharmacol* 232 (2), 309–316, 10.1016/j.taap.2008.07.001. [PubMed: 18671994]
- Oshida K, Vasani N, Jones C, Moore T, Hester S, Nesnow S, Auerbach S, Geter DR, Aleksunes LM, Thomas RS, Applegate D, Klaassen CD, Corton JC, 2015 Identification of chemical modulators of the constitutive activated receptor (CAR) in a gene expression compendium. *Nucl. Recept. Signal* 13 (April), e002, 10.1621/nrs.13002, eCollection 2015. [PubMed: 25949234]
- Oshida K, Vasani N, Thomas RS, Applegate D, Rosen M, Abbott B, Lau C, Guo G, Aleksunes LM, Klaassen C, Corton JC, 2015b Identification of modulators of the nuclear receptor peroxisome proliferator-activated receptor α (PPAR α) in a mouse liver gene expression compendium. *PLOS ONE* 10 (February (2)), e012655, 10.1371/journal.pone.012655, eCollection 2015.
- Oyadomari S, Harding HP, Zhang Y, Oyadomari M, Ron D, 2008 Dephosphorylation of translation initiation factor 2 α enhances glucose tolerance and attenuates hepatosteatosis in mice. *Cell Metab.* 7 (June (6)), 520–532, 10.1016/j.cmet.2008.04.011, [PubMed: 18522833]
- Pery AR, Schuurmann G, Cifroy P, Faust M, Backhaus T, Aicher L, Mombelli E, Tebby C, Cronin MT, Tissot S, Andres S, Brignon JM, Frewer L, Georgiou S, Mattas K, Vergnaud JC, Peijnenburg W, Capri E, Marchis A, Wilks MF, 2013 Perspectives for integrating human and environmental risk assessment and synergies with socio-economic analysis. *Sci. Total Environ* 456–457, 307–316, <http://dx.doi.org/10.1016/j.scitotenv.2013.03.099>.
- Pohjanvirta R, Miettinen H, Sankari S, Hegde N, Linden J, 2012 Unexpected gender difference in sensitivity to the acute toxicity of dioxin in mice. *Toxicol. Appl. Pharmacol* 262 (2), 167–176, <http://dx.doi.org/10.1016/j.taap.2012.04.032>. [PubMed: 22564538]
- Raetzsch CF, Brooks NL, Alderman JM, Moore KS, Hosick PA, Klebanov S, Akira S, Bear JE, Baldwin AS, Mackman N, Combs TP, 2009 Lipopolysaccharide inhibition of glucose production through the Toll-like receptor 4, myeloid differentiation factor 88, and nuclear factor kappa B pathway. *Hepatology* 50 (August (2)), 592–600.
- Ross PK, Woods CG, Bradford BU, Kosyk O, Gatti DM, Cunningham ML, Rusyn I, 2009 Time-course comparison of xenobiotic activators of CAR and PPAR α in mouse liver. *Toxicol. Appl. Pharmacol* 235 (2), 199–207, 10.1016/j.taap.2008.12.011. [PubMed: 19136022]
- Safe S, Lee SO, Jin UH, 2013 Role of the aryl hydrocarbon receptor in carcinogenesis and potential as a drug target. *Toxicol. Sei* 135 (1), 1–16, 10.1093/toxsci/kft128.

- Sand S, Fletcher N, von Rosen D, Kalantari F, Viluksela M, Tuomisto JT, Tuomisto J, Falk-Filipsson A, Hakansson H, 2010 Quantitative and statistical analysis of differences in sensitivity between Long-Evans and Han/Wistar rats following long-term exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Regul. Toxicol. Pharmacol* 57 (2–3), 136–145, <http://dx.doi.org/10.1016/j.yrtph.2010.01.006>. [PubMed: 20138101]
- Sartor MA, Schneckenger M, Marlowe JL, Reichard JF, Wang Y, Fan Y, Ma C, Karyala S, Halbleib D, Liu X, Medvedovic M, Puga A, 2009 Genomewide analysis of aryl hydrocarbon receptor binding targets reveals an extensive array of gene clusters that control morphogenetic and developmental programs. *Environ. Health Perspect* 117 (7), 1139–1146, 10.1289/ehp.0800485. [PubMed: 19654925]
- Schaap MM, Zwart EP, Wackers PF, Huijskens I, van de Water B, Breit TM, van Steeg H, Jonker MJ, Luijten M, 2012 Dissecting modes of action of nongenotoxic carcinogens in primary mouse hepatocytes. *Arch. Toxicol* 86 (11), 1717–1727, 10.1007/s00204-012-0883-6. [PubMed: 22710402]
- Schmidt JV, Bradfield CA, 1996 Ah receptor signaling pathways. *Annu. Rev. Cell Dev. Biol* 12, 55–89, <http://dx.doi.org/10.1146/annurev.cellbio.12.1.55>. [PubMed: 8970722]
- Shaw PJ, Ditewig AC, Waring JF, Liguori MJ, Blomme EA, Ganey PE, Roth RA, 2009 Coexposure of mice to trovafloxacin and lipopolysaccharide, a model of idiosyncratic hepatotoxicity, results in a unique gene expression profile and interferon gamma-dependent liver injury. *Toxicol. Sci* 107 (January (1)), 270–280.
- Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y, 2007 Analysis of gene expression data using BRB-ArrayTools. *Cancer Inf.* 3 (February), 11–17,
- Sobrado M, Pereira MP, Ballesteros I, Hurtado O, Fernández-López D, Pradillo JM, Caso JR, Vivancos J, Nombela F, Serena J, Lizasoain I, Moro MA, 2009 Synthesis of lipoxin A4 by 5-lipoxygenase mediates PPARgamma-dependent, neuroprotective effects of rosiglitazone in experimental stroke. *J. Neurosci* 29 (March (12)), 3875–3884, 10.1523/JNEUROSCI.5529-08.2009. [PubMed: 19321784]
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP, 2005 Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A* 102, 15545–15550.
- Terashima JI, Tachikawa C, Kudo K, Habano W, Ozawa S, 2013 An aryl hydrocarbon receptor induces VEGF expression through ATF4 under glucose deprivation in HepG2. *BMC Mol. Biol* 14 (December), 27. [PubMed: 24330582]
- Tijet N, Boutros PC, Moffat ID, Okey AB, Tuomisto J, Pohjanvirta R, 2006 Aryl hydrocarbon receptor regulates distinct dioxin-dependent and dioxin-independent gene batteries. *Mol. Pharmacol* 69 (1), 140–153, 10.1124/mol.105.018705. [PubMed: 16214954]
- van Kesteren PC, Zwart PE, Pennings JL, Gottschalk WH, Kleinjans JC, van Delft JH, van Steeg H, Luijten M, 2011 Deregulation of cancer-related pathways in primary hepatocytes derived from DNA repair-deficient Xpa^{-/-}p53^{+/-} mice upon exposure to benzo[a]pyrene. *Toxicol. Sci* 123 (September (1)), 123–132, 10.1093/toxsci/kfr169, Epub 2011 Jun 29.
- Ward WO, Delker DA, Hester SD, Thai SF, Wolf DC, Allen JW, Nesnow S, 2006 Transcriptional profiles in liver from mice treated with hepatotumorigenic and nonhepatotumorigenic triazole conazole fungicides: propiconazole, triadimefon, and myclobutanil. *Toxicol. Pathol* 34 (7), 863–878, [PubMed: 17178688]
- Watt K, Jess TJ, Kelly SM, Price NC, McEwan IJ, 2005 Induced alpha-helix structure in the aryl hydrocarbon receptor transactivation domain modulates protein-protein interactions. *Biochemistry* 44 (2), 734–743, 10.1021/bi0487701. [PubMed: 15641800]
- Wilson CL, Miller CJ, 2005 Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. *Bioinformatics* 21 (18), 3683–3685, 10.1093/bioinformatics/bti605. [PubMed: 16076888]

Fig. 1.

AhR biomarker development/characterization and screening of a mouse liver gene expression compendium. Left, biomarker development and characterization. In experiment 1, hepatocytes derived from wild-type and AhR-null mice were treated with benzo[a]pyrene (BaP) or TCDD (Sartor et al., 2009). In experiment 2, wild-type and AhR mice were treated with TCDD (Tijet et al., 2006). Differentially expressed genes (DEGs) were identified using Rosetta Resolver as indicated. Biomarker genes were identified from the DEGs after applying a number of filtering steps. Genes in the biomarker were evaluated by Ingenuity Pathway Analysis (IPA) for canonical pathway enrichment as well as potential transcription factor regulators and by the Comparative Toxicogenomics Database (CTD) to evaluate literature evidence for consistent regulation of biomarker genes by AhR activators. Right, screening and biomarker testing. The AhR biomarker was imported into the NextBio environment in which internal protocols rank ordered the genes based on their fold-change. Screening was carried out by comparison of the biomarker to each bioset in the NextBio database using a pair-wise rank-based enrichment analysis (the Running Fishers test). The

results of the test including the direction of correlation and p -value of the test for each bioset in the compendium were exported and used to populate a master table containing bioset experimental details. A test of the accuracy of the biomarker predictions was carried out with treatments that are known positives and negatives for AhR activation. Screening “hits” were characterized. Additionally, an external gene expression database of experiments performed with Affymetrix gene chips was used in a classification analysis by BRB Array Tools and principal components analysis (PCA). Part of the figure was adapted from a figure in Kupersmidt et al. (2010) and is similar to that in Oshida et al. (2015a,b).

Fig. 2.

Construction and characterization of an AhR biomarker. (A) Expression behavior of genes in the biomarker. The heat map shows the expression of the 90 probe sets after exposure to benzo[a]pyrene *in vitro* (B), TCDD *in vitro* (T1) and TCDD *in vivo* (T2) in liver or primary hepatocytes from wild-type (WT) and *Ahr*-null mice compared to the derived biomarker. (B) Expression behavior of the AhR biomarker genes in the Comparative Toxicogenomics Database (CTD). Four known activators of AhR were examined including TCDD, benzo[a]pyrene (B[a]P), methylcholanthrene (MC), and β -naphthoflavone (β NF). Expression is shown as increased (yellow) or decreased (blue). If two or more references indicate both increased and decreased expression, then the gene is represented as green. The intensity of the genes in the biomarker reflects the fold-change, whereas the intensity of the genes in the CTD analysis reflects the number of publications describing the gene to be altered. (C) Top canonical pathways represented by the genes in the AhR biomarker. Genes were examined by Ingenuity Pathways Analysis. (D) Top transcription factors predicted to

regulate the genes in the AhR biomarker as determined by Ingenuity Pathways Analysis.
(For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Fig. 3.

Prediction of AhR activation using the Running Fisher's test. (A) Heat map showing the relationships between the significance of the positive or negative correlation of contrasts (left and right, respectively) predicted by the Running Fisher's test with the AhR biomarker. Contrasts derived from Affymetrix 430A or 430 2.0 arrays were ordered based on their $-\log(p\text{-value})$. The heat map shows the expression changes (fold-changes) of the corresponding probe sets in the biomarker. The red vertical lines denote the position of biosets with a $p\text{-value} = 10^{-4}$. (B) Comparison of the AhR biomarker to the biosets from wild-type and *Ahr*-*Ahr* mouse tissues (liver or primary hepatocytes) used to derive the biomarker. All $p\text{-values}$ were converted to $-\log 10$ values. Those comparisons which exhibited negative correlation to the biomarker were given a negative value. (C) Biomarker significance following dose- and time-dependent activation of AhR by TCDD. (Left) Three strains of mice were administered TCDD at three dose levels and gene expression was

examined after 24 h (data from E-MEXP-1231). (Middle) Hepatocytes from wild-type mice were treated with 1 nM of TCDD and gene expression was examined at the indicated times (data from Mathijs et al., 2009). (Right) Hepatocytes from wild-type mice or Ahr-null mice expressing the human AhR (hAhR) were treated with 10 nM of TCDD for 6 h (data from Flaveny et al., 2010). (D) Biomarker significance following dose-dependent activation of AhR by benzo[a]pyrene (B[a]P). Hepatocytes from wild-type mice were exposed to four dose levels of B[a]P for 24 h (data from van Kesteren et al., 2011). (E) The AhR biomarker correctly identifies the two known AhR activators (β -naphthoflavone, TCDD) in mice exposed to 12 diverse treatments. See Methods for details of the animal treatments. (F) The biomarker genes separate known positives and negatives for AhR activation using principal components analysis. The first three principal components are shown, derived from the unfiltered expression changes of the biomarker genes across the biosets. Red and green, the three chemical treatments in wild-type or *Ahr*-null mouse tissues used to derive the biomarker, respectively. Blue and purple, biosets positive for AhR activation from *in vitro* and *in vivo* comparisons, respectively. Black, biosets negative for AhR activation. (G) Summary of the sensitivity and specificity of the AhR biomarker. The biomarker was compared to chemicals or other biosets that were known positives or negatives for AhR activation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

64 biosets with AhR activation

25 biosets with AhR suppression

Fig. 4.

AhR activation or suppression in a mouse liver compendium. The AhR biomarker was compared to ~1850 biosets using the Running Fisher's test. The number of biosets with a p -value $< 10^{-4}$ for either activation or suppression in the indicated categories are shown.

Fig. 5.

Chemical activation or suppression of AhR. (A) Distribution of chemical effects on AhR activation or suppression. The p -values for similarity to the AhR biomarker for the 461 chemical comparisons are shown. All p -values were converted to $-\log$ values as described in Fig. 2B. The cutoff values are shown. (B) Chemical activation of AhR. Chemicals which induced AhR ($p\text{-value} = 10^{-4}$) include the known activators dimethylbenzanthracene (DMBA) and aroclor 1254 from the indicated studies. A number of novel activators exhibited significant AhR activation including acetaminophen (APAP; 300 mg/kg for 3 h), AP20189 (0.2 mg/kg for 8 h), GW4064 (3 doses of 150 mg/kg over 24h in liver-specific *Fxr*-null mice), GW501516 (10mg/kg for 8d), lipopolysaccharide (LPS; 2×10^6 EU for 3 h) (Shaw et al., 2009), phenformin (1500 μ M *in vitro* for 48h), quercetin (0.5% in diet for 2 wks), rosiglitazone (10 mg/kg for 8 d), 70 nm silicon dioxide nanoparticles (40 mg/kg for 6 h), and WY (400 μ l of 10 mg/ml in PPAR α -null mice for 6h). (C) Suppression of AhR by CAR activators. (Left) Phenobarbital and TCPOBOP were given to wild-type and Car-null mice for 3 days (GSE40120). (Right) Phenobarbital was given to male and female mice for 2 or 7 days at the indicated dose levels (GSE54597). (D) Effects of PXR activators on AhR. (Left) Wild-type and *Pxr*-null mice were treated with PCN as detailed in the Methods or by "Compound 13" (C13) (from GSE23780). (Right) Activation of AhR in *Pxr*-null mice. Two contrasts comparing *Pxr*-null mice to the corresponding wild-type mice are shown (from GSE55746 and GSE23780, respectively).

Fig. 6. Genetic activation or suppression of AhR. (Left) Effect of the nullizygous mutation in the *Ahr* gene on AhR modulation. Gene expression biosets of *Ahr*-null to wild-type tissue comparisons are shown. (Right) Nullizygous mouse models that affect AhR. The biosets were derived from the following studies: Jun-null (GSE6850), Ghrhr-null (E-MEXP-153), Taf10-null (E-MEXP-3442), Hnf1a-null (E-MEXP-1709), and gp130-null (GSE22009).

Fig. 7. Modulation of AhR by diverse factors. (A) Summary of factors that activate AhR. (B) Summary of factors that suppress AhR. See text for explanation.